IDENTIFIZIERBARKEIT IN MULTIVARIATEN FEHLER-IN-DEN-VARIABLEN-MODELLEN

Inaugural-Dissertation
zur
Erlangung des Grades eines Doktors
der
Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu
BONN

vorgelegt von
Diplom-Mathematiker
HELGE AUFM KAMPE
AUS
LONG BRANCH

BONN 1981

1980

IDENTIFIZIERBARKEIT IN MULTIVARIATEN FEHLER-IN-DEN-VARIABLEN-MODELLEN

Inaugural-Dissertation
zur
Erlangung des Grades eines Doktors
der
Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu
BONN

vorgelegt von
Diplom-Mathematiker
HELGE AUFM KAMPE
aus
LONG BRANCH

BONN 1981

Dekan: Prof. Dr. H.H. Jakobs

Erstberichterstatter: Prof. Dr. P. Schönfeld
Zweitberichterstatter: Prof. Dr. D. Sondermann

Tag der mündlichen Prüfung: 27. Oktober 1980

UNIV.
LEIDEN
BIBL.

Institut für Ökonometrie und Operations Research
Abteilung: Ökonometrie
Direktor: Prof. Dr. P. Schönfeld

Druck: D. Plump, Rheinbreitbach

THEOREM 1.5. *The ring* $\mathrm{Hom}(V, V)$ *is isomorphic to the ring of all n-by-n matrices with elements in k. The isomorphism depends on the choice of a basis. Let g be an element of* $\mathrm{Hom}(V, V)$ *which carries a selected new basis into the old one and suppose that* $g \rightarrow D$ *describes g in terms of the old basis. If* $f \rightarrow A$ *is the description of any f in terms of the old basis, then* DAD^{-1} *is the description of this same f in terms of the new basis.*

Mathematical education is still suffering from the enthusiams which the discovery of this isomorphism has aroused. The result has been that geometry was eliminated and replaced by computations.

[S. 13 aus E.Artin[1957]:Geometric Algebra, Interscience Publ., New York]

VORWORT

Die vorliegende Arbeit wird durch ihre vorwiegend geometrische Argumentation geprägt. Sie fällt somit etwas aus dem Rahmen der traditionellen theoretisch orientierten ökonometrischen Literatur, in der, wie ich meine, von der Geometrie zu wenig Gebrauch gemacht wird.

Während der Abfassung dieser Arbeit habe ich viel von dem an der Abteilung Ökonometrie herrschenden und durch Prof. Dr. P. Schönfeld geprägten wissenschaftlichen Klima profitiert.

Prof. Dr. Schönfeld, Prof. Dr. Deistler, sowie besonders Dr. Hermann gilt mein herzlichster Dank für die Zeit, die sie für zahlreiche Diskussionen mit mir aufgewendet haben.

Die Ausführungen des Abschnitts 3.1.2 gehen auf eine Idee von Dr. Hermann zurück.

INHALTSVERZEICHNIS

0. Einleitung 5-6

1. Eingang und Überblick 7-15
   1.1 Die Modellgleichungen
   1.2 Zwei Beispiele
   1.3 Überblick

2. Identifizierbarkeitskonzepte und konsistente Schätzbarkeit 16-20

3. Verteilungsgestützte Identifizierbarkeit (bzw. Identifizierbarkeit i.s.S.) 21-53
   3.1 Identifikation i.s.S. im strukturellen Fall
      3.1.1 Identifizierbarkeit von Π: Erste Ergebnisse
      3.1.2 Identifizierbarkeit von Π: Ein 'konstruktiver' Beweis des Reiersøl-Resultats
      3.1.3 Identifizierbarkeit von Π: Ein geometrischer Ansatz für den allgemeinen Fall
      3.1.4 Identifizierbarkeit des Rests der Struktur
   3.2 Identifikation i.s.S. im funktionalen Fall

4. Momentengestützte Identifizierbarkeit 54-69
   4.1 Grad-r-momentengestützte Identifizierbarkeit, $r>2$
   4.2 Identifizierbarkeit i.w.S.
      4.2.1 Identifikation bei Nullrestriktionen
      4.2.2 Identifikation bei seriell korrelierten exogenen Variablen

5. Trajektoriengestützte Identifizierbarkeit und konsistente Schätzbarkeit: struktureller vs. funktionaler Fall 70-74
   5.1 Trajektoriengestützte Identifizierbarkeit im strukturellen Fall
   5.2 Konsistente Schätzbarkeit im strukturellen Fall
   5.3 Funktionaler Fall

Literaturverzeichnis 75-80

0. Einleitung

J. Marschak beginnt seine Einführung zum berühmten Cowles Commission Band 10 mit dem Satz: "Quantitative economic study has a threefold basis: it is necessary to formulate economic hypotheses, to collect appropriate data, and to confront hypotheses with data."

Als Ergebnis der ersten Aufgabe, nämlich ökonomische Hypothesen zu formulieren, präsentiert sich in der Regel ein System von Beziehungen zwischen theoretischen Größen, denen zumindest gedanklich eine Meßvorschrift zugeordnet wird.

Die zur Erfüllung der zweiten Aufgabe durchgeführte Erhebung von Daten entspricht sicherlich selten dieser Meßvorschrift. Dies hat zur Folge, daß die tatsächlichen Daten fast immer Realisationen anderer als der 'wahren' theoretischen Größen sind. Auf die Häufigkeit und Schwere dieser Tatsache hat z.B. O. Morgenstern [1965] eindringlich hingewiesen.

Diese Diskrepanz zwischen den 'wahren' und 'beobachtbaren' Variablen wirkt sich selbstverständlich erschwerend auf die Durchführung der dritten der oben genannten Aufgaben aus. T. Haavelmo [1944] formuliert die Situation besonders deutlich: "One could perhaps also characterize the difference between the 'true' and the 'observational' variables in the following way. The 'true' variables are variables such that, if their behaviour should contradict a theory, the theory would be rejected as false; while 'observational' variables, when contradicting the theory, have the possibility that we might be trying out the theory on facts for which the theory was not meant to hold ...

In order to test a theory against facts, or to use it for predictions, either the statistical observations available have to be 'corrected' or the theory itself has to be adjusted, so as to make the facts we consider the

'true' variables relevant to the theory, as described above."

Der Ökonometriker hat meist keinen Einfluß auf die Datenerhebung und kann sich nun z.B. damit behelfen, daß er unbeobachtbare stochastische Störgrößen als Puffer zwischen den 'wahren' und den 'beobachtbaren' Variablen einführt. Diese sogenannten 'Fehler-in-den-Variablen' sind zu unterscheiden von den 'Fehlern-in-den-Gleichungen', welche einzelnen Gleichungen zugeordnet werden, um die Effekte aufzunehmen, die einzeln betrachtet bezüglich der betreffenden Relation vernachlässigbar sind. Wir wollen auf diese Unterscheidung und die dahinter stehenden Überlegungen nicht näher eingehen, da dies bereits anderweitig [z.B. D. Sargan [1958]] geschehen ist.

Werden nun in die meist endlich parametrisierten Relationen Fehler-in-den-Variablen (FV) eingeführt, so ist vor der Schätzung der Parameter - dies ist ja traditionell ein wesentlicher Schritt bei der Gegenüberstellung von Theorie und Daten - zu klären, ob man überhaupt prinzipiell aus den Daten auf die Parameter eindeutig schließen kann. Denn "statistical inference for any number of observations can relate only to characteristics of the observed variables. The limit of statistical inference is an exact knowledge of this distribution function, a limit not attainable but approachable if very large samples can be taken. Anything not implied in this distribution is not a possible object of statistical inference." (Koopmans & Reiersøl [1950])

Dieses 'Identifikationsproblem', dem diese Arbeit im wesentlichen im Rahmen von statischen Modellen gewidmet ist, erlangt gerade in Modellen mit FV große Bedeutung. Die Schwierigkeit, das Problem operational zu lösen, ist der Grund, warum FV-Modelle in der ökonometrischen Literatur immer vernachlässigt wurden§.

§ vergl. Z. Griliches [1974]

1. Eingang und Überblick

1.1 Die Modellgleichungen

Die in der Einleitung angesprochenen theoretischen Beziehungen mögen als lineares Gleichungssystem

$$\Gamma\eta(t) + B\xi(t) = \varepsilon^{(1)}(t) \quad , \quad \operatorname{rg}\Gamma = m \tag{1.1}$$

darstellbar sein. Dabei seien Γ und B Parametermatrizen und t Element einer diskreten Indexmenge $\mathcal{T}$. Die Zufallsvektoren $\eta(t), \operatorname{im}\eta \subset \mathbb{R}^m$, $\xi(t), \operatorname{im}\xi \subset \mathbb{R}^{n-m}$, bzw. $\varepsilon^{(1)}(t)$, $\operatorname{im}\varepsilon^{(1)} \subset \mathbb{R}^m$ enthalten die 'wahren' endogenen Variablen, die 'wahren' exogenen Variablen bzw. die Fehler-in-den-Gleichungen, wobei im· die Bildmenge, d.h. die Menge der Realisationen, des an Stelle des Punktes stehenden Zufallsvektors bezeichnet.

Die Verbindung zu den entsprechenden beobachtbaren endogenen Variablen y(t) und exogenen Variablen x(t) wird durch Einführung der Fehler $\varepsilon^{(2)}(t)$ und $\delta(t)$ hergestellt:

$$\begin{aligned} x(t) &= \xi(t) + \delta(t) \\ y(t) &= \eta(t) + \varepsilon^{(2)}(t). \end{aligned} \tag{1.2}$$

Es erweist sich als sinnvoll, die Gleichungen (1.1) und (1.2) in die 'finale' Form zu bringen und dabei noch einige neue Bezeichnungen einzuführen:

$$\begin{aligned} z(t) := \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \begin{pmatrix} I \\ -\Gamma^{-1}B \end{pmatrix} \xi(t) + \begin{pmatrix} \delta(t) \\ \Gamma^{-1}\varepsilon^{(1)}(t) + \varepsilon^{(2)}(t) \end{pmatrix} \\ &=: \begin{pmatrix} I \\ \Pi \end{pmatrix} \xi(t) + \begin{pmatrix} \delta(t) \\ \varepsilon(t) \end{pmatrix} \\ &=: \Lambda\xi(t) + \nu(t) \\ &=: \zeta(t) + \nu(t) \end{aligned}$$

Man spricht von 'struktureller' Spezifikation, wenn die Variablen $\xi(t)$ für alle $t \in \mathcal{T}$ unabhängig identisch verteilte Zufallsvektoren sind und von 'funktionaler' Spezifikation, wenn sie vollkommen degeneriert, also deter-

ministisch sind.

Die Störvariablen $\nu(t)$, die sich unabhängig von $\zeta(t)$ realisieren mögen, werden als unabhängig identisch - meist normal - verteilt angenommen.

Die Grundform

$$z(t) = \Lambda \xi(t) + \nu(t) \qquad (1.3)$$

tritt außer in den Wirtschaftswissenschaften auch in anderen empirischen Wissenschaften auf.

So ist (1.3) Ausgangspunkt der Faktoranalyse, die hauptsächlich in der Psychometrie Anwendung findet. Dort heißen die in $\xi(t)$ auftretenden Variablen gemeinsame Faktoren und jene in $\nu(t)$ spezielle Faktoren. Die Matrix Λ hat aber im allgemeinen nicht die spezielle Struktur des FV-Modells.

In der Soziometrie wird vielfach mit Pfeildiagrammen gearbeitet, in denen Kausalpfade zwischen soziologischen Größen dargestellt werden, die formal in einem System wie (1.3) zusammengefaßt werden können. Entsprechend ihrem Einfluß bekommen die Pfade als Gewichte die sogenannten Pfadkoeffizienten zugeordnet. Diese errechnen sich dadurch, daß man in (1.3) zu standardisierten Größen übergeht, zweite Momente bildet und die enstandenen Gleichungen löst.§

1.2 Zwei Beispiele

Bevor wir einen Überblick über die Ergebnisse dieser Arbeit und deren Einordnung in die bestehende Literatur geben, wollen wir zwei Beispiele vorführen. Sie dienen einem dreifachen Zweck. Zum ersten erfüllen sie die Form (1.3) mit etwas Leben, zum zweiten führen sie in die mit (1.3) verbundene Identifikationsproblematik ein, und zum dritten machen sie den Leser etwas mit dem später verwen-

§ vergl. hierzu z.B. O.D.Duncan [1975], Kap. 2. Die Aussage ist nicht ganz streng zu nehmen. Wie wir in Abschnitt sehen werden, treten hier Lösbarkeitsprobleme auf.

deten Instrumentarium vertraut.

Die Stochastik von (1.3) in struktureller Spezifikation können wir in einem gewissen Umfang durch die Darstellung der zweiten Momente mit Hilfe von Konzentrationsellipsoiden veranschaulichen.

Malinvaud [1970](S.154) definiert das Konzentrationsellipsoid eines n-dimensionalen Zufallsvektors w mit Kovarianzmatrix Ω als die Menge C der n-dimensionalen Vektoren c, für die gilt:

$$|\gamma' c| \leq 1 \text{ für alle } \gamma \text{ mit } \gamma'\Omega\gamma \leq 1 .$$

Ist Ω nicht singulär, so ist die Beschreibung des Konzentrationsellipsoids als die Menge der Vektoren c, für die $c'\Omega^{-1}c \leq 1$ gilt, bequemer. Um eine entsprechende Beschreibung für singuläres Ω zu erhalten, führen wir den Begriff des Trägers eines Zufallsvektors ein. Dieser Begriff wird auch weiter hinten noch häufiger auftreten. Der Träger ca(imw) [carrier $\hat{=}$ (engl.)Träger] eines Zufallsvektors w sei die kleinste lineare Mannigfaltigkeit, die imw mit Wahrscheinlichkeit 1 (W 1) enthält. Hat nun Ω den Rang $r < n$ und ist der Träger von w gegeben durch $R'w = 0$ mit einer $n\times(n-r)$ Matrix R, die den Rang n-r hat und $\Omega R = 0$ erfüllt, so ist das Konzentrationsellipsoid die Menge der Vektoren c, für die gleichzeitig $R'c = 0$ und $c'(\Omega+RR')^{-1}c \leq 1$ gilt(siehe Malinvaud [1970] S.156f).

Konkret mögen jetzt für die beiden folgenden Beispiele in (1.3) nur eine exogene und eine endogene Variable auftreten. Die Beziehung zwischen diesen sei im ersten Beispiel gegeben durch $\eta = \xi$ - also $\pi = 1$ - und die Kovarianzmatrizen von ν bzw.ζ seien

$$\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} .$$

Die entsprechenden Konzentrationsellipsoide sind gegeben durch:

$$C_\nu = \left\{ \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \middle| \; \nu_1^2 + 2\nu_1\nu_2 + 2\nu_2^2 \leq 1 \right\}$$

und

$$C_\zeta = \left\{ \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \middle| \; \zeta_1^2 + \zeta_2^2 \leq 2, \; \zeta_1 = \zeta_2 \right\} .$$

Beobachtbar ist aber nur z bzw. hier nach Annahme dessen Konzentrationsellipsoid

$$C_z = \{\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mid \frac{z_1^2}{3} + \frac{z_2^2}{2} \leq 1\},$$

welches sich aus der Kovarianzmatrix $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ von z herleiten läßt.

Andererseits kann dieses Konzentrationsellipsoid aber auch unter der Annahme $\pi=0$ aus den Ellipsoiden

$$C_{\tilde{\nu}} = \{\begin{pmatrix} \tilde{\nu}_1 \\ \tilde{\nu}_2 \end{pmatrix} \mid \tilde{\nu}_1^2 + \frac{\tilde{\nu}_2^2}{2} \leq 1\}$$

und

$$C_{\tilde{\zeta}} = \{\begin{pmatrix} \tilde{\zeta}_1 \\ \tilde{\zeta}_2 \end{pmatrix} \mid \frac{\tilde{\zeta}_1^2}{2} \leq 1,\ \tilde{\zeta}_2 = 0\}$$

hervorgegangen sein, wie sofort zu sehen ist, wenn man die zugehörigen Kovarianzmatrizen $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ von $\tilde{\nu}$ und $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ von $\tilde{\zeta}$ addiert.

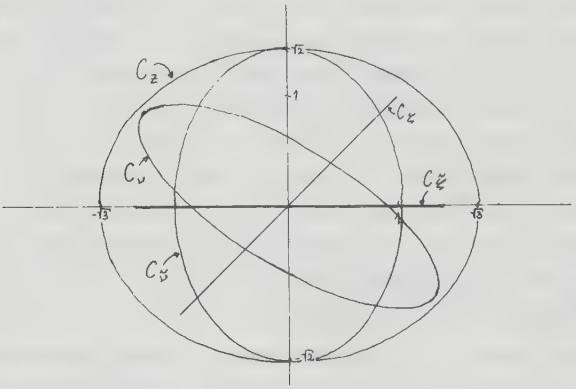


Abb. 1

Dieses von Malinvaud [1970](S.401) stammende Beispiel zeigt, daß mindestens zwei unterschiedliche Festlegungen zum gleichen beobachtbaren Ellipsoid führen. Sie sind somit an Hand der Information, die in diesen Konzentrationsellipsoiden steckt, nicht unterscheidbar, d.h. nicht identifizierbar. Anders ausgedrückt, ist die wahre Struktur dieses Modells aus zweiten Momenten nicht identifizierbar.

Wir wollen ein zweites Beispiel betrachten, bei dem wir nun aber die Verteilungen selbst benutzen. Hier nehme

ξ nur zwei Werte an, sagen wir 0 und 1, jeweils mit der Wahrscheinlichkeit 1/2. ζ nimmt infolgedessen die Werte $\begin{pmatrix}0\\0\end{pmatrix}$ und $\begin{pmatrix}1\\\pi\end{pmatrix}$ jeweils mit Wahrscheinlichkeit 1/2 an. Beobachtbar ist die Faltung dieser diskreten Verteilung mit der angenommenen Normalverteilung von ν.

In diesem Fall ließe sich π durch folgende heuristische Überlegung identifizieren. Die beobachtete Verteilung wird sich immer durch Verunreinigungen von der Normalverteilung unterscheiden, es sei denn man betrachtet sie in Richtung der Geraden, die durch die beiden Punkte geht, in denen sich ζ realisiert. Das bedeutet, daß π dadurch bestimmt ist, daß die Projektion nur längs des Vektors $(1,\pi)'$ normal ist.

Dies können wir präzisieren, indem wir an Stelle der Verteilungen charakteristische Funktionen (c.F.) benutzen.§ Die c.F. $\phi_\zeta(t)$ von ζ hat mit den obigen Annahmen die folgende Gestalt:

$$\phi_\zeta(t) = \phi_{\begin{pmatrix}\zeta_1\\\zeta_2\end{pmatrix}}\begin{pmatrix}t_1\\t_2\end{pmatrix} = \phi_{\begin{pmatrix}1\\\pi\end{pmatrix}\xi}\begin{pmatrix}t_1\\t_2\end{pmatrix} = \phi_\xi\left[(1,\pi)\begin{pmatrix}t_1\\t_2\end{pmatrix}\right]$$

$$= \phi_\xi(t_1+\pi t_2) = \frac{1}{2} + \frac{1}{2}\exp\{i(t_1+\pi t_2)\}.$$

Die c.F. von ν sei der Einfachheit halber

$$\phi_\nu(t) = \phi_{\begin{pmatrix}\nu_1\\\nu_2\end{pmatrix}}\begin{pmatrix}t_1\\t_2\end{pmatrix} = \exp\{-\frac{1}{2}(\sigma_1^2t_1^2 + \sigma_2^2t_2^2)\}.$$

Dann ist die c.F. von z:

$$\phi_z(t) = \phi_\zeta(t)\cdot\phi_\nu(t)$$

$$= \frac{1}{2}\exp\{-\frac{1}{2}(\sigma_1^2t_1^2 + \sigma_2^2t_2^2)\}$$

$$+ \frac{1}{2}\exp\{i(t_1+\pi t_2) - \frac{1}{2}(\sigma_1^2t_1^2 + \sigma_2^2t_2^2)\}.$$

Die c.F. von z nach einer Projektion, die durch die idempotente Matrix P beschrieben werden soll, lautet dann allgemein:

$$\phi_{Pz}(t) = \phi_\zeta(P't)\cdot\phi_\nu(P't).$$

---

§ Für die Begründung dieses Vorgehens und für die Definition der c.F. vergl. Abschnitt 3.1.1

Solange $\phi_{\zeta}(P't)$ von t abhängt, kann $\phi_{Pz}(t)$ wegen des Satzes von Cramér(vergl. Theorem 4.6.3 in Cuppens[1975]) nicht die c.F. einer Normalverteilung sein. $\phi_{\zeta}(P't)$ wird aber nur dann unabhängig von t sein, wenn für alle t

$$(1,\pi)'P't = 0$$

gilt. Daraus folgt aber

$$P\begin{pmatrix}1\\ \pi\end{pmatrix} = 0 \ ,$$

was nichts anderes heißt als, daß P entlang $(1,\pi)'$ projiziert.

1.3 Überblick

Um das gerade an speziellen Beispielen aufgezeigte Identifikationsproblem allgemein vernünftig angehen zu können, führen wir im zweiten Abschnitt unterschiedliche Identifizierbarkeits- und Schätzbarkeitskonzepte ein. Sie sind das entscheidende Instrument zur Strukturierung der Probleme in FV-Modellen.

So liefert die Unterscheidung zwischen verteilungs-, momenten- und trajektoriengestützter Identifizierbarkeit einerseits und konsistenter Schätzbarkeit andererseits eine Gliederung, unter der die Identifizierbarkeitsbedingungen und Schätzmethoden für FV-Modelle sinnvoll geordnet werden können und das Verhältnis des strukturellen Falls zum funktionalen Fall deutlich wird: Während bei funktionaler Spezifikation die Koeffizienten zwar verteilungs- bzw. momentengestützt identifizierbar, aber ohne weitere Annahmen nicht trajektoriengestützt identifizierbar bzw. konsistent schätzbar sind, ist der Sachverhalt bei struktureller Spezifikation umgekehrt. Hier stellt die Sicherung der verteilungs- oder momentengestützten Identifizierberkeit ein Problem dar. Liegt diese vor, so ist, wie in Abschnitt 5 dargelegt wird, die trajektoriengestützte Identifizier-

barkeit und konsistente Schätzbarkeit nicht mehr problematisch.

Im dritten Abschnitt über verteilungsgestützte Identifizierbarkeit finden sich die wesentlichsten neuen Beiträge. Hier wird das bekannte Resultat von Reiersøl[1950], daß im Falle einer endogenen und einer exogenen Variablen die Identifizierbarkeit des Koeffizienten der exogenen Variable genau dann gewährleistet ist, wenn diese Variable nicht normalverteilt ist, auf den Fall mehrerer Variablen verallgemeinert. Dabei treten Schwierigkeiten auf, da der Begriff der Nichtnormalität im mehrdimensionalen Fall zunächst nicht eindeutig ist und verschiedene Präzisierungen zuläßt.

Als erstes benutzen wir eine naheliegende, aber starke Forderung der Nichtnormalität an ξ, nämlich die, daß keine Linearkombination der Komponenten von ξ normalverteilt sein soll. Nichtnormalität in diesem Sinne ist eine hinreichende Bedingung zur Identifizierbarkeit von Π. Dies hat bereits J. Wolfowitz[1952] gezeigt, jedoch auf einem anderen Weg, als wir ihn beschreiten. Unser Vorgehen ermöglicht es, den Einfluß von a-priori-Information über Π einzubeziehen.

Auf der Suche nach einer passenden Definition, die eine notwendige und hinreichende Bedingung liefert, greifen wir in 3.1.2 noch einmal den Reiersøl-Fall auf und zeigen, wie π 'konstruktiv' zu identifizieren ist. Die Idee besteht darin, aus den c.F. der beobachtbaren Variablen normale Faktoren mit maximaler Varianz abzuspalten und aus den verbleibenden 'nichtnormalen Kernen' π zu bestimmen. Diese Idee ließ sich nicht auf den allgemeinen Fall ausdehnen, da dieses algebraische Verfahren sehr direkt ist und die Eindimensionalität ausnützt. Trotzdem gewährt sie Einsicht und motiviert eine zweite allgemeine Definition der Nichtnormalität, die auf T.A. Jeeves[1954] zurückgeht und auf weniger direkten geometrischen Überlegungen beruht.

Er definiert den Grad der Nichtnormalität von z als

das Minimum der Dimensionen der linearen Unterräume L mit der Eigenschaft, daß sich z als Summe zweier unabhängiger Zufallsvariablen u und v darstellen läßt, wobei v normalverteilt und L Träger von u ist. Ist die Nichtnormalität von z nun gleich dem Rang der Matrix Λ aus (1.3), schreibt Jeeves, kann man den Raum, der von den Spalten von Λ aufgespannt wird, identifizieren. In seiner Argumentation steckt aber ein Fehler, der seine Beweisidee unbrauchbar zu machen scheint. Deshalb müssen wir nach einem vollkommen anderen Weg suchen. Dies geschieht im Abschnitt 3.1.3. Zum Schluß jenes Abschnitts geben wir noch eine dritte Definition der Nichtnormalität, die ebenfalls von Jeeves stammt. Sie rührt von einem weiteren Ansatz 'konstruktiven' Identifizierens her und ist Grundlage einer hinreichenden Bedingung, deren Notwendigkeit noch in Frage steht.

Auf die Identifizierbarkeit weiterer Elemente der Struktur gehen wir im Abschnitt 3.1.4 in einfacher Erweiterung der Ergebnisse von Reiersøl ein.

Unseres Wissens hat sich außer Jeeves nur noch C.R. Rao[1966],[1967],[1969] mit Verallgemeinerungen von Reiersøls Resultaten befaßt. Seine Annahmen sind aber wesentlich stärker als die unsrigen. Er setzt voraus, daß die Komponeneten von ξ und von ν jeweils untereinander unabhängig sind. Dies versetzt ihn in die Lage, unter Umgehung vieler Probleme, die durch unsere Annahmen entstehen, mit Funktionalgleichungen und ihren Lösungen zu operieren und daraus etwas stärkere Resultate zu erzielen, die in A.M. Kagan,Y.V. Linnik,C.R. Rao[1973], Kap. 10 zusammengefaßt sind.

Im vierten Hauptabschnitt schließlich beschäftigen wir uns mit momentengestützter Identifizierbarkeit. In 4.1 führen wir eine Idee von Geary[1942] multivariat aus. Es werden Beziehungen zwischen Kumulanten von z und den Elementen von Π abgeleitet, aus denen unter Umständen die Koeffizienten eindeutig bestimmt werden können.

In 4.2 beschränken wir uns auf Momente zweiter Ord-

nung und zeigen, wie man mit der Methode der Instrumentenvariablen Bedingungen für Identifizierbarkeit gewinnen kann. Die erste Anwendung bildet in 4.2.1 der Fall, wo Nullrestriktionen für strukturelle Parameter vorliegen. Hier ergeben sich notwendige Bedingungen, die auf andere Weise bereits von V. Geraci[1976] und C. Hsiao[1976] gefunden wurden. Eine zweite Anwendung des Instrumentenvariablengedankens läßt sich bei zeitlich abhängigen Beobachtungen demonstrieren.

2. Identifizierbarkeitskonzepte und konsistente Schätzbarkeit

Das Identifikationsproblem und damit auch der Begriff der Identifizierbarkeit haben in den fundamentalen Arbeiten von L. Hurwicz[1950] und T.C. Koopmans & O. Reiersøl [1950] eine verbindliche Formulierung erfahren. Später wurden durch T.A. Jeeves[1954] und H. Bunke & O. Bunke [1974] noch leichte Verallgemeinerungen vorgenommen. Die Verbindung mit konsistenter Schätzbarkeit ist erst in letzter Zeit durch die Arbeiten von P. Gänssler[1969], P. Schönfeld[1975] und M. Deistler & H.-G. Seifert[1978] wieder in den Vordergrund gerückt worden, nachdem offenbar bereits früher ein Interesse daran bestanden hat, wie aus den Arbeiten von J. Wolfowitz[1957], L. LeCam[1958] und L. Breiman, L. LeCam & L. Schwartz[1964] zu schließen ist.

Zur Definition der Identifizierbarkeit benötigen wir noch einige vorbereitende Definitionen und Bemerkungen. Zunächst sei vereinbart, daß alle Zufallsvariablen auf einem zugrundeliegenden Wahrscheinlichkeitsraum $(E,\mathcal{E},P)$ definiert sind.

Definition 2.1 (Koopmans&Reiersøl)

Unter einer S t r u k t u r $\sigma = (F_\omega, \Phi)$ verstehen wir eine genau spezifizierte Wahrscheinlichkeitsverteilung F_ω von latenten stochastischen Variablen ω und eine genau spezifizierte strukturelle Beziehung

$$\Phi(w,\omega) = 0$$

zwischen beobachtbaren Variablen w und den latenten Variablen ω, die die eindeutige Bestimmung der beobachtbaren Variablen aus den latenten Variablen mit W 1 erlaubt.

Als Beispiel diene (1.3). Hier ist $w := z$; $\omega := \begin{pmatrix}\xi\\ \nu\end{pmatrix}$ und $\Phi(w,\omega) := (I,-\Lambda,-I)\begin{pmatrix}z\\ \xi\\ \nu\end{pmatrix}$.

Definition 2.2

Ein ökonometrisches Modell $\mathfrak{M}$ ist der Inbegriff aller bezüglich des studierten ökonomischen Phänomens a priori getroffenen Hypothesen.

Das Modell stellt sich demnach als ein Katalog von Modellannahmen dar und läßt sich als die Menge Σ der unter den Annahmen zulässigen Strukturen σ auffassen. Als Bilder von Σ unter surjektiven Abbildungen s_K und $s_{\mathcal{F}}$ unterscheiden wir die Menge K von Charakteristiken κ der strukturellen Beziehung Φ und der latenten Variablen ω und die Menge $\mathcal{F}$ der durch σ's induzierten Verteilungen F_w der beobachtbaren Variablen.

Jedes σ legt eindeutig eine Verteilung $F_w = s_{\mathcal{F}}(\sigma)$ fest. Identifizierbarkeit hängt mit der Umkehrung dieser Aussage zusammen. Für eine allgemeine Definition der Identifizierbarkeit führen wir auf K und $\mathcal{F}$ Äquivalenzrelationen $\overset{K}{\sim}$ und $\overset{\mathcal{F}}{\sim}$ ein, die jeweils in konkreten Fällen durch die Probleme verschiedene Spezifikationen erfahren. Die mit $\overset{K}{\sim}$ und $\overset{\mathcal{F}}{\sim}$ erzeugten Äquivalenzklassen indizieren wir mit $k \in K$ und $\ell \in L$ und nennen sie interne bzw. externe Information. Neben $\overset{K}{\sim}$ wird durch $\overset{\mathcal{F}}{\sim}$ auf K eine weitere Äquivalenzbeziehung, die sogenannte Beobachtungsäquivalenz, induziert: $\kappa \overset{\mathcal{F}}{\sim} \tilde{\kappa} :\Leftrightarrow s_{\mathcal{F}}(\sigma) \overset{\mathcal{F}}{\sim} s_{\mathcal{F}}(\tilde{\sigma})$ und $\kappa = s_K(\sigma), \tilde{\kappa} = s_K(\tilde{\sigma})$.

Definition 2.3 (Bunke&Bunke)

Die Charakteristik κ heißt $(\overset{K}{\sim}, \overset{\mathcal{F}}{\sim})$ - identifizierbar in $\mathfrak{M}$ $\Leftrightarrow$ für beliebige Paare $(\kappa, F), (\tilde{\kappa}, \tilde{F}) \in K \times \mathcal{F}$ gilt: $F \overset{\mathcal{F}}{\sim} \tilde{F} \Rightarrow \kappa \overset{K}{\sim} \tilde{\kappa}$.

Bemerkung 2.4 (vergl. Schönfeld [1975], proposition 1)

κ $(\overset{K}{\sim}, \overset{\mathcal{F}}{\sim})$ - identifizierbar in $\mathfrak{M}$ $\Leftrightarrow$ es existiert eine Funktion $\chi_\Sigma : L \to K$, so daß $k = \chi_\Sigma(\ell)$, d.h. man kann aus

der externen Information eindeutig auf die interne Information schließen <==> die externe Information bezüglich κ ist feiner als die interne Information.

Wir erhalten verschiedene Identifizierbarkeitskonzepte, wenn wir $\mathcal{C}$ auf verschiedene Weisen festlegen:
(i) Wenn $F \mathcal{C} \tilde{F}$:<==> $F = \tilde{F}$, werden wir von v e r t e i l u n g s g e s t ü t z t e r I d e n t i f i z i e r b a r k e i t bzw. von Identifizierbarkeit im strengen Sinne (i.s.S.) sprechen. Hier besteht die externe Information aus der Kenntnis der Verteilungsfunktion.
(ii) $M_r(F) = M_r(\tilde{F})$ bedeute, daß die Verteilungen F und $\tilde{F}$ in ihren Momenten bis zur r-ten Ordnung übereinstimmen, wobei die Existenz der Momente vorausgesetzt wird. Wenn dann $F \mathcal{C} \tilde{F}$:<==> $M_r(F) = M_r(\tilde{F})$, werden wir von G r a d r - m o m e n t e n g e s t ü t z t e r I d e n t i f i z i e r b a r k e i t bzw. speziell im Fall r = 2 von Identifizierbarkeit im weiteren Sinne (i.w.S.) reden.

Bemerkung 2.5

Für zwei Äquivalenzbeziehungen $\mathcal{C}_1$, $\mathcal{C}_2$, die in der Beziehung

$$F \mathcal{C}_1 \tilde{F} \Rightarrow F \mathcal{C}_2 \tilde{F} \text{ für alle } F, \tilde{F} \in \mathcal{F}$$

stehen, gilt:

$$\kappa\ (\mathcal{K}, \mathcal{C}_2)\text{-identifizierbar} \Rightarrow \kappa\ (\mathcal{K}, \mathcal{C}_1)\text{-identifizierbar.}$$

Gilt die Aussage der Definition 2.3 nur für $\tilde{\kappa}$'s aus einer offenen Umgebung von κ, so heißt κ lokal $(\mathcal{K}, \mathcal{C})$-identifizierbar.

Die bisherigen Identifizierbarkeitsbegriffe gehen von Kenntnissen über den gesamten Wahrscheinlichkeitsraum $(E, \mathcal{E}, P)$ aus, also etwa von der Kenntnis der Verteilung F oder deren theoretischer Momente. Im Hinblick auf die praktische Schätzsituation ist ein Identifizierbarkeitsbegriff, der auf Trajektorien aufbaut, wünschenswert.

Hiermit beschäftigt sich ausführlich die Arbeit von Deistler & Seifert [1978]. Wir geben hier nur eine knappe

Darstellung. Obgleich unsere Bezeichnungen etwas von denen bei Deistler&Seifert abweichen, ist die Übertragung klar, so daßes keine Schwierigkeiten bereiten wird, für die Beweise zu den Bemerkungen den angegebenen Hinweisen nachzugehen.

Wir führen auf $\mathcal{K}$, K, und L Borelsche σ-Algebren $\mathcal{B}_K$, $\mathcal{B}_K$, und $\mathcal{B}_L$ ein und betrachten eine Trajektorie $W^{\mathcal{J}}(e) := (w(1,e), w(2,e), \ldots)$, $e \in E$ des beobachtbaren n-dimensionalen Zufallsvektors $w(t,\cdot)$, die als Punkt im $(\mathbb{R}^{n\mathcal{J}}, \mathcal{B}^{\mathcal{J}})$ interpretiert wird. Gegebenenfalls sei noch ein Maß P_K auf $\mathcal{B}_K$ gegeben. $P_w^{(\kappa)}$ sei das durch die Charakteristik κ der Struktur σ bestimmte Maß über $\mathcal{B}^{\mathcal{J}}$.

Definition 2.6 (LeCam)

Die interne (bzw. externe) Information k (bzw. ℓ) heißt [fast überall] e r r e i c h b a r in $\mathcal{M}$ <===> es existiert eine $(\mathcal{B}^{\mathcal{J}}, \mathcal{B}_K$ (bzw. $\mathcal{B}_L))$ - meßbare Funktion

$$\hat{k} \text{ (bzw. } \hat{\ell}): \mathbb{R}^{n\mathcal{J}} \to K \text{(bzw. L)},$$

so daß $P_w^{(\kappa)}\{W^{\mathcal{J}}(e) \mid \hat{k}(W^{\mathcal{J}}(e)) = k \text{ (bzw. } \hat{\ell}(W^{\mathcal{J}}(e)) = \ell)\} = 1$ für [P_K - fast] alle $\kappa \in \mathcal{K}$.

Definition 2.7

Die Charakteristik κ heißt [f.ü.] t r a j e k t o r i e n g e s t ü t z t i d e n t i f i z i e r b a r <===> κ ist [f.ü.] erreichbare interne Information oder $(\mathcal{B}_L, \mathcal{B}_K)$ - meßbare Funktion einer [f.ü.] erreichbaren externen Information.

Statt von Erreichbarkeit bzw. von trajektoriengestützter Identifizierbarkeit sprechen Gänssler, Schönfeld und Deistler&Seifert von 'exakter Schätzbarkeit'. Wir haben trajektoriengestützte Identifizierbarkeit so wie oben definiert, um den Unterschied zwischen zwei grundsätzlich möglichen Schätzansätzen zu verdeutlichen. Der Teil bis zum 'oder' in der Definition 2.7 würde bereits genügen,

denn es gilt die

Bemerkung 2.8 (vergl. Deistler&Seifert, Beweis zu theorem 5)

κ Funktion einer erreichbaren externen Information ==> κ ist erreichbare interne Information.

Die Schätzsituation wird beschrieben durch eine aufsteigende Folge von σ-Algebren $\mathcal{B}^T \subset \mathcal{B}^{T+1} \subset \ldots \subset \mathcal{B}^{\infty}$, wobei $\mathcal{B}^{\infty}$ von $\bigcup_{T=1}^{\infty} \mathcal{B}^T$ erzeugt wird, und eine Folge von $(\mathcal{B}^T, \mathcal{B}_K)$-meßbaren Funktionen $\hat{\kappa}_T: \mathbb{R}^{nT} \to K$.

Definition 2.9

Die Charakteristik κ heißt [f.ü.] s t a r k k o n s i s t e n t s c h ä t z b a r in $\mathcal{M}$ <===> es existiert eine Folge von $(\mathcal{B}^T, \mathcal{B}_K)$-meßbaren Funktionen $\hat{\kappa}_T$, so daß $P_w^{(\kappa)} \{w^{\infty}(e) \mid \lim_{T\to\infty} \hat{\kappa}_T(w^{\infty}(e)) = \kappa\} = 1$ für [P_K-fast] alle $\kappa \in K$.

Bemerkung 2.10 (vergl. Deistler&Seifert)

κ [f.ü.] stark konsistent schätzbar ==> κ [f.ü.] trajektoriengestützt identifizierbar;
κ [f.ü.] trajektoriengestützt identifizierbar ==> κ fast überall stark konsistent schätzbar

Im fünften Abschnitt wird auch von schwach konsistenten Schätzern bzw. Schätzerfolgen die Rede sein. Dies bedeutet dann, daß in Definition 2.9 statt der Konvergenz mit Wahrscheinlichkeit eins Konvergenz nach Wahrscheinlichkeit vorliegt.

3. Verteilungsgestützte Identifizierbarkeit (bzw. Identifizierbarkeit i.s.S.)

Wir beginnen die Untersuchungen zur Identifizierbarkeit mit dem Maximalfall verfügbarer Information, d.h. die externe Information bestehe in diesem dritten Hauptabschnitt aus der Kenntnis der Verteilungsfunktionen der beobachtbaren Variablen $z(t)$. Die Äquivalenzrelation auf $\mathcal{F}$ sei also durch

$$F_{z(t)} \overset{\mathcal{F}}{\sim} \tilde{F}_{z(t)} : \Longleftrightarrow F_{z(t)} = \tilde{F}_{z(t)}$$

gegeben.

3.1 Identifikation i.s.S. im strukturellen Fall

Wir gehen hier vom System (1.3), also von

$$z(t) = \begin{pmatrix} I \\ \Pi \end{pmatrix} \xi(t) + \nu(t)$$

aus und treffen zunächst die Annahmen:

(A) $(\xi(t))$ sei eine Folge von unabhängig identisch verteilten $(n-m)$-dimensionalen Zufallsvektoren

(B) $(\nu(t))$ sei eine Folge von unabhängig identisch verteilten n-dimensionalen Zufallsvektoren. Der Prozeß $(\nu(t))$ sei stochastisch unabhängig vom Prozeß $(\xi(t))$.

Die Annahme (A) ist kennzeichnend für den 'strukturellen' Fall. Zusammen mit Annahme (B) hat sie zur Folge, daß wir bei Untersuchungen zur Identifizierbarkeit den Parameter t fallen lassen können. Wir weisen daraufhin, daß wir hier keine Annahmen bezüglich der Existenz von Momenten machen.

3.1.1 Identifizierbarkeit von Π: Erste Ergebnisse

Unser Hauptanliegen ist es, Aussagen über die Identifizierbarkeit von Π zu gewinnen. Dazu wollen wir uns zu-

nächst ansehen, welche Folgerungen bezüglich Π wir aus der Beobachtungsäquivalenz zweier Strukturen $\sigma,\tilde{\sigma} \in \Sigma$, $\Sigma = \{(\Pi,F_\xi,F_\nu)\,|\,(1.3),(A),(B)\}$ erhalten. Wir setzen also $K = \{\Pi\,|\,(\Pi,F_\xi,F_\nu) \in \Sigma\}$ und für $\overset{*}{\sim}$ die Gleichheit.

Die Beobachtungsäquivalenz drücken wir statt mit Verteilungen mit den zugehörigen c.F. aus. Diese Methode, die Reiersøl in seiner bekannten Arbeit[1950] benutzt, ist durch die multivariate Fassung des Eindeutigkeitssatzes für c.F. gerechtfertigt.

Satz 3.1

Zwei Verteilungsfunktionen F_z und $\tilde{F}_z$ sind genau dann gleich, wenn ihre c.F.

$$\phi_z(t) := \int \exp(it'\lambda)\,dF_z(\lambda)$$

und

$$\tilde{\phi}_z(t) := \int \exp(it'\lambda)\,d\tilde{F}_z(\lambda)$$

gleich sind.

Beweis: (siehe z.B. Cuppens[1975], Theorem 2.2.2) ▲

Da nun

$$\phi_z(t) = \phi_{\binom{x}{y}}\binom{t_1}{t_2} = E\{\exp(it_1'x + it_2'y)\}$$

$$= E\{\exp[it_1'(\xi+\delta) + it_2'(\Pi\xi+\varepsilon)]\}$$

$$= E\{\exp[i(t_1+\Pi' t_2)'\xi]\}E\{\exp(it_1'\delta+it_2'\varepsilon)\}$$

$$= \phi_\xi(t_1+\Pi' t_2)\phi_{\binom{\delta}{\varepsilon}}\binom{t_1}{t_2},$$

gilt:

$$\phi_z(t)=\tilde{\phi}_z(t) \iff \phi_\xi(t_1+\Pi' t_2)\phi_{\binom{\delta}{\varepsilon}}\binom{t_1}{t_2} = \tilde{\phi}_\xi(t_1+\tilde{\Pi}' t_2)\tilde{\phi}_{\binom{\delta}{\varepsilon}}\binom{t_1}{t_2}. \quad (3.1)$$

Durch Setzen von $t_1 = -\tilde{\Pi}' t_2$ folgt:

$$\phi_\xi((\Pi-\tilde{\Pi})' t_2)\cdot\phi_{-\tilde{\Pi}\delta+\varepsilon}(t_2) = \tilde{\phi}_{-\tilde{\Pi}\delta+\varepsilon}(t_2)$$

bzw.

$$\phi_{(\Pi-\tilde{\Pi})\xi}(t_2)\cdot\phi_{-\tilde{\Pi}\delta+\varepsilon}(t_2) = \tilde{\phi}_{-\tilde{\Pi}\delta+\varepsilon}(t_2) \quad (3.2)$$

für alle Vektoren $t_2 \in \mathbb{R}^m$, speziell also für Vektoren, die

außer an der i-ten Stelle, i=1,...,m, nur Nullen haben und dessen i-te Komponente alle reelle Zahlen durchlaufen kann. Dann folgt aus (3.2)

$$\phi_{(\Pi_{i\cdot}-\tilde{\Pi}_{i\cdot})\xi}(t)\cdot\phi_{-\tilde{\Pi}_{i\cdot}\delta+\varepsilon_i}(t)=\phi_{-\tilde{\Pi}_{i\cdot}\delta+\varepsilon_i}(t) \qquad (3.3)$$

für i=1,...,m, wobei $\Pi_{i\cdot}$ die i-te Zeile von Π bezeichnet. Wir fassen das Ergebnis zusammen in

Satz 3.2

Gegeben (1.3),(A),(B), so gilt für i=1,...,m :

$$\Pi_{i\cdot} \overset{\mathcal{K}}{\in} \tilde{\Pi}_{i\cdot} \quad \Longrightarrow \quad (\Pi_{i\cdot}-\tilde{\Pi}_{i\cdot}) \text{ erfüllen (3.3).}$$

Dies liefert bereits eine Identifizierbarkeitsaussage:

Satz 3.3

Im Modell [(1.3),(A),(B)] ist Π i.s.S. identifizierbar, wenn es keine (nichttriviale) Linearkombination der Komponenten von ξ gibt, so daß die c.F. hiervon Faktor der c.F. einer Linearkombination der Fehler ist.

Beweis: Aus $\Pi_{i\cdot} \overset{\mathcal{K}}{\in} \tilde{\Pi}_{i\cdot}$ folgt wegen Satz 3.2 , daß $(\Pi_{i\cdot}-\tilde{\Pi}_{i\cdot})$ (3.3) erfüllt. Die Bedingung des Satzes impliziert aber $\Pi_{i\cdot}-\tilde{\Pi}_{i\cdot}=0$, also $\Pi_{i\cdot} \overset{\mathcal{K}}{\in} \tilde{\Pi}_{i\cdot}$. ▲

Diese Aussage ist noch sehr allgemein, da sie von den Verteilungsannahmen über ξ und ν abhängt. Wir wollen nun die klassische Normalverteilungsannahme über den Fehler ν machen. Um auch degenerierte Normalverteilungen zu erfassen, treffen wir folgende Vereinbarung:

(C) Der n×1-Vektor ν habe die Darstellung

$$\nu = D\Delta,$$

wobei der stochastische d×1-Vektor Δ eine zentrierte normale Dichte

$$f_{\Delta}(t) = (2\pi)^{-\frac{1}{2}d}\exp(-\tfrac{1}{2}t't)$$

mit $t \in \mathbb{R}^d$ besitzt und D eine $n \times d$- Matrix ist.

Das so definierte ν hat den Erwartungswert O und eine Kovarianzmatrix $\Omega := DD'$, die singulär sein kann. Gelten zu (1.3) die Annahmen (A), (B) und (C), so werden wir künftig vom Modell $\mathfrak{M}_1$ sprechen.

Wir betrachten nun alle Linearkombinationen der Komponenten von ξ, sagen wir $g'\xi$, die (eindimensional) normalverteilt sind mit Erwartungswert O und fassen die Zeilenvektoren g' in der Menge $\mathfrak{G}$ zusammen. Da nach dem Satz von Cramér die c.F. einer Normalverteilung nur normale Faktoren besitzt, ist $\mathfrak{G}$ auch die Menge aller Vektoren g', für die g'x normalverteilt ist. Man beachte, daß $\mathfrak{G}$ kein linearer Raum ist, da die Additivität verletzt sein kann.

Mit dem Satz von Cramèr erhalten wir eine Verschärfung des Satzes 3.2.

Satz 3.4

In $\mathfrak{M}_1$ gilt: $\Pi_{i.} \overset{\mathcal{L}}{\sim} \Pi_{i.} \implies \Pi_{i.} - \Pi_{i.} \in \mathfrak{G}$,i=1,...,m .

Daraus folgt sofort

Satz 3.5

$\mathfrak{G} \equiv \{O\}$ ==> für alle i=1,...,m ist
$\Pi_{i.}$ $(\mathcal{L}, \mathcal{E})$ - identifizierbar in $\mathfrak{M}_1$.

Dieses Ergebnis, daß $\Pi_{i.}$ verteilungsgestützt identifizierbar ist, wenn keine (nichttriviale) Linearkombination der Komponenten von ξ zentriert normalverteilt ist, hat bereits J. Wolfowitz[1952] auf andere Weise erhalten.

Bemerkung 3.6

Sind alle Komponenten von $\Pi_{i.}$ bis auf eine, sagen wir π_{ij_0}, identifizierbar und ist $E\xi_{j_0} \neq 0$, so ist auch π_{ij_0} iden-

tifizierbar. Denn aus (3.3) und der Annahme (C) geht hervor, daß $(\pi_{ij_0} - \tilde{\pi}_{ij_0})\xi_{j_0}$ normalverteilt sein muß mit Erwartungswert 0, wenn π_{ij_0} und $\tilde{\pi}_{ij_0}$ beobachtungsäquivalent sind. Das impliziert aber mit der gemachten Voraussetzung, daß $\pi_{ij_0} = \tilde{\pi}_{ij_0}$ ist. Speziell sind also im Fall von Absolutgliedern diese identifizierbar, wenn die übrigen Koeffizienten identifizierbar sind.

Für weitere Ergebnisse betrachten wir den linearen Raum, der von den Zeilenvektoren in $\mathcal{H}$ aufgespannt wird und den Orthogonalraum hierzu bezüglich $\mathbb{R}^{n-m}$. Der erstgenannte Raum sei mit $\langle\mathcal{H}\rangle$ bezeichnet und der letzere mit $\langle\mathcal{H}\rangle^{\perp}$. In $\langle\mathcal{H}\rangle^{\perp}$ wählen wir eine Basis aus Spaltenvektoren $\bar{g}_j$, $j=1,\ldots,n-m-\dim\langle\mathcal{H}\rangle$, die in der Matrix $\bar{G}$ zusammengefaßt seien.

Satz 3.7

In $\mathfrak{M}_1$ gilt für alle $i=1,\ldots,m$:

$$\Pi_{i\cdot} \,\mathcal{R}\, \tilde{\Pi}_{i\cdot} \quad ==> \quad \Pi_{i\cdot}\bar{G} = \tilde{\Pi}_{i\cdot}\bar{G} \ .$$

Beweis: $\Pi_{i\cdot} \,\mathcal{R}\, \tilde{\Pi}_{i\cdot}$ => $(\Pi_{i\cdot} - \tilde{\Pi}_{i\cdot}) \in \mathcal{H}$ => $(\Pi_{i\cdot} - \tilde{\Pi}_{i\cdot}) \in \langle\mathcal{H}\rangle$

<==> für alle $\bar{g} \in \langle\mathcal{H}\rangle^{\perp}$: $(\Pi_{i\cdot} - \tilde{\Pi}_{i\cdot})\bar{g} = 0$

<==> $(\Pi_{i\cdot} - \tilde{\Pi}_{i\cdot})\bar{G} = 0$. ▲

Hieraus folgt sofort der

Satz 3.8

In $\mathfrak{M}_1$ gilt für alle $\bar{g} \in \langle\mathcal{H}\rangle^{\perp}$:

$\Pi_{i\cdot}\bar{g}$ ist i.s.S. identifizierbar.

Wir wollen nun das Wissen um solch eine identifizierende Abbildung $\Pi_{i\cdot}' \to \bar{G}'\Pi_{i\cdot}'$ ausnutzen, um eine Bedingung dafür anzugeben, wann eine reellwertige Funktion $\mathcal{A}$ von $\Pi_{i\cdot}'$ unter zusätzlichen Restriktionen lokal identifizierbar ist. Die zusätzlichen Restriktionen seien

in Form einer vektorwertigen Funktion $\mathcal{K}: \mathbb{R}^{n-m} \to \mathbb{R}^{k}$ von $\Pi'_{i.}$ gegeben, deren Werte bekannt sein mögen. Wir nehmen an, daß sowohl h als auch $\mathcal{K}$ einmal stetig differenzierbar sind und kürzen dies mit $h, \mathcal{K} \in \mathcal{C}^1$ ab. Mit $\mathcal{J}_{\mathcal{K}}(\Pi'^{0}_{i.})$ und $\operatorname{grad} h(\Pi'^{0}_{i.})$ seien die k×(n-m) Jakobi-Matrix von $\mathcal{K}$ und der 1×(n-m) Gradient von h jeweils an der Stelle $\Pi'^{0}_{i.}$ bezeichnet. Der 'if'-Teil des Theorems 2.1 von J.Richmond[1976]§ liefert uns

Satz 3.9

Seien $h, \mathcal{K} \in \mathcal{C}^1$. Dann ist $h(\Pi'_{i.})$ in $\mathcal{M}_1$ lokal identifizierbar in einer Umgebung von $\Pi'^{0}_{i.}$, wenn

$$\operatorname{rg}(\bar{G}: \mathcal{J}_{\mathcal{K}}(\Pi'^{0}_{i.}): \operatorname{grad}' h(\Pi'^{0}_{i.})) = \operatorname{rg}(\bar{G}: \mathcal{J}_{\mathcal{K}}(\Pi'^{0}_{i.})) \qquad (3.4)$$

gilt.

Beweis: Wegen Satz 3.8 ist $\Pi\bar{G}$ identifizierbar. Dann ist $h(\Pi'_{i.})$ nach Bemerkung 2.4 lokal identifizierbar, wenn es eine Funktion f gibt, so daß $h(\Pi'_{i.}) = f(\bar{G}'\Pi'_{i.}, \mathcal{K}(\Pi'_{i.}))$ in einer Umgebung von $\Pi'^{0}_{i.}$, da $(\bar{G}'\Pi'_{i.}, \mathcal{K}(\Pi'_{i.}))$ eine Funktion der externen Information bezüglich $\Pi'_{i.}$ ist. Theorem 2.1 von Richmond schließt den Beweis. ▲

---

§ (Richmond[1976]): 'Throughout this section we shall deal with two functions $F, K \in \mathcal{C}^1$, that is F and K are continuous with continuous first order partial derivatives. It will be assumed that $F: \mathbb{R}^n \to \mathbb{R}^m$ and for simplicity that $K: \mathbb{R}^n \to \mathbb{R}$; we are thus dealing with an (m+1)-tuple of functions $(F,K) = (f_1, \ldots f_m, K)$. We shall denote by $\mathcal{J}_F(\cdot)$ the m×n Jacobian matrix of F and by $\nabla K(\cdot)$ the 1×n gradient vector of K. A point $x_0 \in \mathbb{R}^n$ will be said to be regular with respect to F if $\operatorname{rank}(\mathcal{J}_F(x))$ is constant for every x in some neighborhood of x_0; x_0 is regular with respect to the pair (F,K) if $\operatorname{rank}(\mathcal{J}_F(x) : \nabla' K(x))$ is constant for every x in some neighborhood of x_0.

Theorem 2.1 Let x_0 be regular with respect to F and with respect to the pair (F,K). Then there exists a neighborhood U of x_0 such that there exists a function g such that $K = g \circ F$ over U if and only if

$$\operatorname{rank}(\mathcal{J}_F(x_0) : \nabla' K(x_0)) = \operatorname{rank}(\mathcal{J}_F(x_0)).$$

Bemerkung 3.10

Ist ξ multinormalverteilt, so wissen wir aus dem Abschnitt 1.2, daß Π im allgemeinen nicht identifizierbar ist. In diesem Fall ist $\bar{G} \equiv 0$ und Satz 3.9 zeigt, daß Π identifizierbar ist, wenn geeignete zusätzliche Annahmen über Π gemacht werden können. Wir werden im vierten Abschnitt sehen, wie Nullrestriktionen an die strukturellen Parameter, die im klassischen simultanen System ohne Fehler-in-den-Variablen überidentifizierend wären, also Restriktionen an Π darstellen, dazu dienen, Π i.w.S. zu identifizieren.

Wir weisen noch auf ein einfaches und bekanntes Ergebnis hin:

Bemerkung 3.11

Ist in $\mathfrak{M}_1$ Ω bekannt und $\xi \neq 0$, so ist Π i.s.S. identifizierbar, denn (3.1) liefert

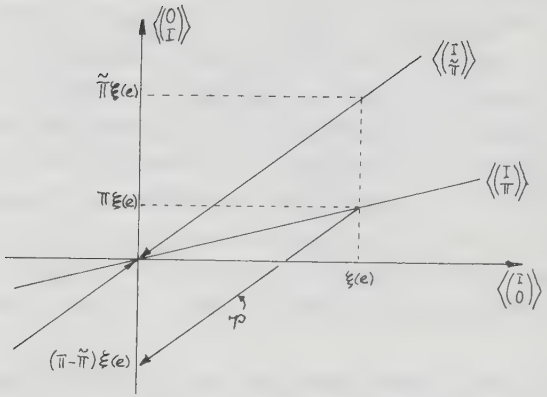
$$\phi_{\xi}[(\Pi-\tilde{\Pi})'t_2] = 1 \text{ für alle } t_2 \in \mathbb{R}^m,$$

woraus $\Pi = \tilde{\Pi}$ folgt, wenn $\xi \neq 0$.

Bisher haben wir nur hinreichende Bedingungen für die Identifizierbarkeit von Π angegeben. Um der Frage nach der Notwendigkeit der Bedingung in Satz 3.5 nachzugehen, wollen wir den Schritt von (3.1) nach (3.2) geometrisch interpretieren.

In (3.1) gehen wir von zwei verschiedenen Unterräumen $\left\langle \begin{smallmatrix} I \\ \Pi \end{smallmatrix} \right\rangle$ und $\left\langle \begin{smallmatrix} I \\ \tilde{\Pi} \end{smallmatrix} \right\rangle$ aus, auf denen sich jeweils ζ und $\tilde{\zeta}$ realisieren. Zu beobachten ist die Verteilung F_z, die sowohl durch Faltung der Verteilung von $\zeta = \begin{pmatrix} I \\ \Pi \end{pmatrix} \xi$ mit einer Normalverteilung, als auch durch Faltung der Verteilung von $\tilde{\zeta} = \begin{pmatrix} I \\ \tilde{\Pi} \end{pmatrix} \xi$ mit einer weiteren Normalverteilung entsteht. Das Setzen von $t_1 = -\tilde{\Pi}' t_2$ entspricht geometrisch der Durchführung einer durch die Matrix $\begin{pmatrix} 0 & 0 \\ -\tilde{\Pi} & I \end{pmatrix}$ beschriebenen Abbildung. Dies ist eine Projektion entlang $\left\langle \begin{smallmatrix} I \\ \tilde{\Pi} \end{smallmatrix} \right\rangle$ auf $\left\langle \begin{smallmatrix} 0 \\ I \end{smallmatrix} \right\rangle$ (siehe Abb.2).

Abb. 2



Nach Durchführung der Projektion sehen wir, daß $(\Pi-\tilde{\Pi})\xi$ multinormalverteilt ist. Gibt es umgekehrt eine Matrix G mit $G\xi$ normalverteilt, so erfüllt $\tilde{\Pi} = \Pi-G$ die für die Beobachtungsäquivalenz notwendige Bedingung, daß $(\Pi-\tilde{\Pi})\xi$ $(= G\xi)$ normalverteilt ist. Um für die Beobachtungsäquivalenz auch hinreichend zu sein, muß $\tilde{\Pi}$ so beschaffen sein, daß es ein ξ und ein davon unabhängiges $\tilde{\nu}$ gibt, so daß $\tilde{\Pi}\xi+\tilde{\nu} = z$ ist. Dafür muß es einen Unterraum geben mit der Eigenschaft, daß die (bedingten) Verteilungen auf allen zu diesem Unterraum parallelen linearen Mannigfaltigkeiten, unabhängig von den (die Mannigfaltigkeiten charakterisierenden) Werten von ξ, normal sind und mit der Verteilung von $\begin{pmatrix}I\\\Pi\end{pmatrix}\xi$ gefaltet F_z ergeben.

Ergebnisse darüber, ob ein solcher Unterraum immer existiert, sind mir nicht bekannt, so daß wir über die Notwendigkeit der Identifizierbarkeitsbedingung in Satz 3.5 nichts Endgültiges sagen können.

3.1.2 Identifizierbarkeit von Π: Ein 'konstruktiver' Beweis des Reiersøl-Resultats

Auf der Suche nach einer hinreichenden und notwendigen Bedingung für den allgemeinen Fall greifen wir noch einmal das einfachere Problem auf, das Reiersøl[1950] unter-

sucht hat und geben einen 'konstruktiven' Beweis seines Resultats, indem wir direkt zeigen, wie die Nichtnormalität von ξ dazu benutzt werden kann, um Π zu identifizieren. Der Beweis gewährt weiteren Einblick in die Struktur des Problems und könnte darüber hinaus als Grundlage eines konsistenten Schätzverfahrens dienen.

Reiersøl betrachtet in seiner Arbeit[1950] die Gleichungen

$$x = \xi + \delta \tag{3.5a}$$

$$y = \pi_0 + \pi_1 \xi + \varepsilon \tag{3.5b}$$

unter den Annahmen (A),(B) und (C). Wir nehmen ferner hier der Einfachheit halber an, daß Erwartungswerte existieren.

Zu unserem alternativen konstruktiven Beweis, wie die Nichtnormalität von ξ die Identifizierbarkeit von π_0 und π_1 sichert, benötigen wir drei einfache Lemmata. Die darin auftretenden Zufallsvariablen mögen alle den Erwartungswert 0 haben.

Lemma 3.12 (Rao[1967] S.1211)

Jede reelle Zufallsvariable w kann in eindeutiger Weise zerlegt werden in

$$w = w_1 + w_2 ,$$

wobei w_1 und w_2 unabhängig sind, w_2 normalverteilt ist und w_1 keinen normalen Faktor außer dem trivialen besitzt. Dabei bedeutet "u ist Faktor von w", daß man die c.F. von w faktorisieren kann in $\phi_u(t) \cdot \phi_{\bar{u}}(t)$, wobei $\phi_u(t)$ und $\phi_{\bar{u}}(t)$ c.F. sind.

Beweis: Mit $\phi_w(t)$ sei die c.F. von w bezeichnet. Betrachten wir nun die Menge aller σ^2, so daß $\phi_w(t) \cdot \exp(\frac{1}{2} \sigma^2 t^2)$ c.F. ist. Diese Menge ist von oben beschränkt, da $|\phi_w(t) \cdot \exp(\frac{1}{2} \sigma^2 t^2)| \leq 1$. Ihr Supremum sei σ^{*2}. Die Funktion $\phi_w(t) \cdot \exp(\frac{1}{2} \sigma^{*2} t^2)$ ist Grenzwert

einer Folge von c.F. und stetig bei t = 0. Damit ist sie nach dem Stetigkeitssatz für c.F. (vergl. corollary 2 zu theorem 3.6.1 in E.Lukacs[1970] S. 51) selbst c.F. Bezeichnen wir sie mit $\phi_{w_1}(t)$ und setzen wir $\phi_{w_2}(t) := \exp(-\frac{1}{2}\sigma^{*2}t^2)$, so erhalten wir das gewünschte Resultat. ▲

w_2 nennen wir die maximale normale Komponente von w und w_1 bezeichnen wir als nichtnormalen Kern von w.

Lemma 3.13

Es gelte $w = \omega + \alpha$, wobei ω und α stochastisch unabhängig sind und α normalverteilt ist. Dann ist der nichtnormale Kern w_1 von w gleich ω_1, dem nichtnormalen Kern von ω.
Beweis: Angenommen w_2 wäre eine normale Komponente mit $\sigma^2_{w_2} > \sigma^2_{\omega_2} + \sigma^2_{\alpha}$. Dann müßte gelten: $w_2 = \omega_2 + \alpha + \mu$, wobei μ normal ist und folglich:

$$\phi_w = \phi_{w_1}\phi_{\omega_2}\phi_\mu\phi_\alpha = \phi_{\omega_1}\phi_{\omega_2}\phi_\alpha \text{ bzw. } \phi_{\omega_1} = \phi_{w_1}\phi_\mu$$

im Gegensatz zur Voraussetzung, daß ω_1 nichtnormaler Kern von ω ist. ▲

Lemma 3.14

Es sei $w = \pi\rho, \pi \in \mathbb{R}\setminus\{0\}$. Der nichtnormale Kern w_1 von w ist $\pi\rho_1$, wobei ρ_1 nichtnormaler Kern von ρ ist.
Beweis: Wir suchen nach

$$\tau^{*2} = \max\{\tau^2 \mid \phi_w(t)\cdot\exp(\tfrac{1}{2}\tau^2t^2) \text{ c.F.}\}$$

(mit $\pi t = s$) $$= \max\{\tau^2 \mid \phi_\rho(s)\cdot\exp(\tfrac{1}{2}(\tau^2/\pi^2)s^2) \text{ c.F.}\}.$$

Das Maximum wird für $\tau^{*2} = \pi^2\sigma^{*2}$ angenommen, so daß die maximale normale Komponente von w wegen
$\phi_{\pi\rho_2}(t) = \exp(-\frac{1}{2}\pi^2\sigma^{*2}t^2)$ $\pi\rho_2$ ist. ▲

Mit den gemachten Annahmen zu (3.5a-b) wissen wir, daß

$$E\xi = Ex \text{ und } E\eta = Ey = \pi_0 + \pi_1 Ex \qquad (3.6)$$

ist und können somit (3.5a-b) überführen in

$$\hat{x} = \xi + \delta$$
$$\hat{y} = \pi_1 \xi + \varepsilon \ ,$$

wobei jetzt alle Variablen den Erwartungswert 0 haben. Wenn ξ nicht normalverteilt ist -wobei immer unter "normalverteilt" auch "degeneriert normalverteilt" (vergl. Annahme (C)) zu verstehen ist- so können wir die nichtnormalen Kerne $\hat{x}_1$ von $\hat{x}$ und $\hat{y}_1$ von $\hat{y}$ bestimmen. Auf Grund der Lemmata 3.12-3.14erhalten wir

$$\hat{y}_1 = (\pi_1 \xi)_1 = \pi_1 \xi_1 = \pi_1 \hat{x}_1 \ .$$

Daraus lassen sich π_1 und mit (3.6) π_0 z.B. dadurch eindeutig bestimmen, daß wir die zugehörigen charakteristischen Funktionen in einem Punkt vergleichen. Sie sind somit i.s.S. identifizierbar. ▲

Die Idee dieses Beweises konnte nicht auf den allgemeinen Fall, d.h. für beliebige n und $m<n$ übertragen werden. Der Satz von Cramér und Wold[1936], mit dem man oft die Untersuchung eines mehrdimensionalen Wahrscheinlichkeitsmaßes auf die Untersuchung seiner eindimensionalen Projektionen zurückführen kann, hilft hier nicht weiter, da das Supremum des normalen Faktors entscheidend von der Richtung im $\mathbb{R}^n$ abhängt. Dabei verstehen wir, R. Cuppens[1975] folgend, unter der Projektion des Maßes P auf einen Vektor (die Richtung) $t_0 \in \mathbb{R}^n$ das Maß $\hat{P}$, welches durch

$$\hat{P}(N) = P\{x \mid x \cdot t_0 \in N\} \text{ für alle Borelmengen } N \subset \mathbb{R}$$

definiert wird.

Ein direktes Maximieren der Kovarianzmatrix des normalen Faktors scheitert daran, daß dies nicht mit der linearen Struktur verträglich ist, d.h. es gibt keine Ordnung, die gegenüber linearen Abbildungen invariant ist.

3.1.3 Identifizierbarkeit von Π: Ein geometrischer Ansatz für den allgemeinen Fall

Das obige algebraische Identifikationsverfahren ist sehr direkt und liefert mehr als eine Identifizierbarkeitsaussage. Deshalb sind die Schwierigkeiten bei der Verallgemeinerung der Methode einerseits nicht überraschend und andererseits kein Indiz dafür, daß keine Identifizierbarkeitsaussagen über die in Abschnitt 3.1.1 hinaus zu erzielen wären. Wir beschreiten nun einen weniger direkten, geometrischen Weg, der auf der Idee in der Arbeit von Jeeves[1954] aufgebaut ist.

Hierzu gehen wir von der geometrischen Interpretation der Gleichungskette (1.3) mit den Annahmen (A)-(C) aus, nämlich, daß z sich in zwei stochastisch unabhängige Komponenten additiv zerlegen läßt, von denen die eine sich im linearen Unterraum, der von den Spalten von Λ aufgespannt wird, realisiert und die andere normalverteilt ist und Werte im ganzen Raum annehmen kann.

In unserem FV-Fall ist (im Gegensatz zum Modell der Faktoranalyse) die Identifizierbarkeit von Λ mit der Identifizierbarkeit des Spaltenraumes $\langle\Lambda\rangle$ von Λ äquivalent. Denn während die Identifizierbarkeit von Λ klarerweise die Identifizierbarkeit von Λ impliziert, überlegt man sich umgekehrt, daß Λ auf Grund der speziellen Gestalt $\begin{pmatrix} I \\ \Pi \end{pmatrix}$ durch $\langle\Lambda\rangle$ eindeutig bestimmt wird. Denn für ein alternatives $\tilde{\Lambda} = \begin{pmatrix} I \\ \tilde{\Pi} \end{pmatrix}$ muß, wegen $\left\langle \begin{smallmatrix} I \\ \Pi \end{smallmatrix} \right\rangle = \left\langle \begin{smallmatrix} I \\ \tilde{\Pi} \end{smallmatrix} \right\rangle$, $\begin{pmatrix} I \\ \Pi \end{pmatrix} = \begin{pmatrix} I \\ \tilde{\Pi} \end{pmatrix} P$ und folglich $\Pi = \tilde{\Pi}$ gelten.

Wir untersuchen also künftig die Identifizierbarkeit von $\langle\Lambda\rangle$, d.h. die Äquivalenzrelation $\overset{*}{\sim}$ ist nun gegeben durch

$$\Lambda \overset{*}{\sim} \tilde{\Lambda} \iff \bigvee_{P} \Lambda = \tilde{\Lambda} P \quad (P \text{ nichtsingulär})$$

Wie wir in 3.1.1 gesehen haben, hängt die Identifizierbarkeit mit der Normalität bzw. Nichtnormalität von ξ bzw. z zusammen. Für mehrdimensionale Variablen ist aber der Begriff der Nichtnormalität zunächst nicht eindeutig.

Die Bedingung von 3.1.1, daß keine Linearkombination der Komponenten von x normalverteilt sein soll, scheint sehr stark. Andererseits wird man einen Zufallsvektor schon dann nicht mehr als normalverteilt bezeichnen, wenn nur eine Komponente nicht normalverteilt ist, was aber, wie wir sehen werden, im allgemeinen zur Identifizierbarkeit von Λ nicht ausreicht.

Wir werden also bei der Definition der Nichtnormalität das Ausmaß der Nichtnormalität genauer berücksichtigen müssen. Diese Überlegungen und jene aus dem vorangehenden Unterabschnitt motivieren die folgende

Definition 3.15 (Jeeves)

Die Nichtnormalität im ersten Sinne $nn_1(w)$ eines n-dimensionalen Zufallsvektors w ist das Minimum der Dimensionen der linearen Unterräume L mit der Eigenschaft, daß sich w als Summe zweier unabhängiger Zufallsvariablen u und v darstellen läßt, wobei v normal und L Träger von u ist.

Wenn w normalverteilt ist, wählen wir für u den Nullvektor und $nn_1(w)$ ist dann O. Wenn andererseits die c.F. von w keinen nichttrivialen normalen Faktor besitzt, setzen wir v = O.

Zur Motivation des folgenden Lemmas stellen wir einige elementare geometrische Überlegungen an. Hierzu erinnern wir an die Definition des Konzentrationsellipsoids aus Abschnitt 1.2 und nehmen zur Illustration an, daß die Matrix Ω nichtsingulär ist. Zwei Vektoren x und y heißen konjugiert bezüglich des Konzentrationsellipsoids, wenn $x'\Omega^{-1}y = O$ gilt. Es sei nun als Beispiel ein normalverteilter zentrierter Zufallsvektor $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ gegeben. Er läßt sich als Summe $\omega_1+\omega_2=w_1\begin{pmatrix}1\\0\end{pmatrix}+w_2\begin{pmatrix}0\\1\end{pmatrix}$ zweier Zufallsvariablen mit orthogonalen Trägern interpretieren. ω_1 und ω_2 sind genau dann unabhängig, wenn $Ew_1w_2=O$, d.h. wenn die Kovarianzmatrix Ω von $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ Diago-

nalgestalt besitzt. Die Träger von ω_1 und ω_2 sind mithin konjugiert bezüglich Ω^{-1}. Transformiert man nun mit einer nichtsingulären Matrix T, so bleiben $T\omega_1$ und $T\omega_2$ stochastisch unabhängig und ihre Träger sind konjugiert bezüglich $T'^{-1}\,\Omega^{-1}T^{-1}$. In Umkehrung dieser Überlegungen können wir durch Transformation immer erreichen, daß sich jede Normalverteilung durch Faltung zweier Zufallsvariablen mit orthogonalen Trägern erzeugen läßt.

Konkret sei die Verteilung charakterisiert durch

$$\Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}, \quad (\Omega \text{ nichtsingulär})$$

und gesucht ist die nichtsinguläre lineare Transformation $T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$, die letzteres erreicht.

Wir wollen den (eindimensionalen) Unterraum $\left\langle\!\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle\!\right\rangle$ invariant lassen und fordern, daß $\det(T) = 1$, so daß T die Form

$$\begin{pmatrix} 1 & t_{12} \\ 0 & 1 \end{pmatrix}$$

haben muß. Die bezüglich des Konzentrationsellipsoids zu $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ konjugierten Vektoren $\begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$ errechnen sich aus

$$0 = [1/\det(\Omega)]\,(k_1, k_2) \begin{pmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = [1/\det(\Omega)]\,(k_1\sigma_2^2 - k_2\sigma_{12})$$

zu $\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} k_2\sigma_{12}/\sigma_2^2 \\ k_2 \end{pmatrix}$.

Nach dem vorher Gesagten suchen wir nun nach dem T, welches $\begin{pmatrix} k_2\sigma_{12}/\sigma_2^2 \\ k_2 \end{pmatrix}$ in $\lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\lambda \in \mathbb{R}$ überführt, bzw. welches

$$\lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & t_{12} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} k_2\sigma_{12}/\sigma_2^2 \\ k_2 \end{pmatrix} = \begin{pmatrix} k_2\sigma_{12}/\sigma_2^2 + k_2 t_{12} \\ k_2 \end{pmatrix}$$

erfüllt. Hieraus ergibt sich

$$T = \begin{pmatrix} 1 & -\sigma_{12}/\sigma_2^2 \\ 0 & 1 \end{pmatrix}.$$

Die durch dieses T beschriebene Abbildung ist eine 'Scherung'. Diese Bezeichnung erklärt sich aus der folgen-

den Zeichnung.

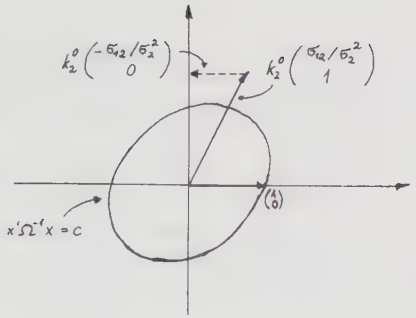


Abb. 3

Allgemein gilt

Lemma 3.16

Sei $nn_1(w) = \ell$ und L der von den ersten ℓ karthesischen Basisvektoren aufgespannte Unterraum. Dann gibt es eine lineare Abbildung (Scherung) $\mathcal{A}$ so, daß $\mathcal{A}w = w_1 + w_2$, wobei $ca(imw_1) = L$, w_1 und w_2 stochastisch unabhängig sind, w_2 normal ist und $ca(imw_2)$ orthogonal ($\perp$) zu L liegt.

Beweis: Zunächst erreichen wir durch eine orthogonale Transformation T (also $T' = T^{-1}$), daß $Tw = Tu + Tv = \begin{pmatrix}u_1\\0\end{pmatrix} + \begin{pmatrix}v_1\\v_2\end{pmatrix}$, wobei $ca(imv_1) \subset ca(imu_1) = L$. Anschließend definieren wir eine Abbildung $\tilde{\mathcal{A}}$ durch $\tilde{\mathcal{A}}:\begin{pmatrix}k_1\\k_2\end{pmatrix} \to \begin{pmatrix}k_1 - Sk_2\\k_2\end{pmatrix}$, S eine Matrix, und erhalten mit

$\tilde{w}_1 := \begin{pmatrix}u_1\\0\end{pmatrix} + \begin{pmatrix}v_1 - Sv_2\\0\end{pmatrix}$ und $\tilde{w}_2 := \begin{pmatrix}0\\v_2\end{pmatrix}$

$\tilde{\mathcal{A}}(Tw) = \tilde{w}_1 + \tilde{w}_2$, wobei ersichtlich $ca(im\tilde{w}_1) \perp ca(im\tilde{w}_2)$. S wird dabei so gewählt, daß $E(v_1 - Sv_2)v_2' = 0$ bzw. $Ev_1v_2' = SEv_2v_2'$ gilt. Ein solches S existiert immer, da für alle Vektoren k, die $(Ev_2v_2')k = 0$ erfüllen, auch $(Ev_1v_2')k = 0$ gilt. Da v zentriert normalverteilt ist, sind $v_1 - Sv_2$ und v_2 stochastisch unabhängig. Da außerdem u und v und somit Tu und Tv stochastisch unabhängig sind, sind $\tilde{w}_1$ und $\tilde{w}_2$ stochastisch unabhängig. Umkehrung der orthogonalen Transformation T liefert die gewünschte Ab-

bildung. Es ist $w_1'w_2 = (T^{-1}\tilde{w}_1)'(T^{-1}\tilde{w}_2) = \tilde{w}_1'(TT^{-1})\tilde{w}_2 = 0$ und $\det(\mathcal{A}) = 1$. ▲

Für die Nichtnormalität im ersten Sinne erhalten wir im multivariaten Fall ähnliche Aussagen wie die der Lemmata 3.13 und 3.14.

Lemma 3.17

Sei $w = \omega + \alpha$, ω und α stochastisch unabhängig und α normalverteilt. Dann gilt: $nn_1(w) \leq nn_1(\omega)$.
Beweis: Sei $nn_1(\omega) = \ell_1$. Dann sollen in $\mathcal{A}_1\omega = \omega_1 + \omega_2$, $\mathcal{A}_1$, ω_1 und ω_2 die in Lemma 3.16 beschriebenen Eigenschaften besitzen. ω_1 und ω_2 lassen sich als Projektionen von $\mathcal{A}_1\omega$ darstellen, so daß sie jeweils von α unabhängig sind. $\omega_2 + \alpha$ ist somit normalverteilt und unabhängig von ω_1, woraus schließlich folgt, daß $nn_1(w) \leq \dim[ca(im\omega_1)] = \ell_1$. ▲

Ob auch $nn_1(\omega) \leq nn_1(w)$ gilt, ist ein offenes Problem. Hierzu wäre auszuschließen, daß durch die Faltung von ω mit α in einer zusätzlichen Dimension Nichtnormalität verschwindet. Das so etwas nicht von vorne herein auszuschließen ist, mag folgendes eindimensionales Beispiel von Lukacs[1970](S.184) andeuten.

$\phi_1(t) = (1 - 1/2\ t^2)\cdot\exp(-t^2/4)$ ist eine unzerlegbare c.F. und somit weder selbst normal noch hat sie einen normalen Faktor. Aber $\phi_2(t) = [\phi_1(t)]^2 = (1-t^2+t^4/4).\exp(-t^2/2$ $= [(1-t^2/2)^2\exp(-3t^2/8)][\exp(-t^2/8)]$ liefert eine Faktorisierung in c.F., bei der ein Faktor normal ist.

Lemma 3.18

Sei $w = \Lambda\rho$, wobei $rg\,\Lambda = \dim[ca(im\rho)]$, w $(n\times1)$-, ρ $[(n-m)\times1]$- und Λ $[n\times(n-m)]$-Matrizen sind. Dann ist $nn_1(w) = nn_1(\rho)$.
Beweis: Sei $nn_1(\rho) = \ell_2$ und $\mathcal{A}_1\rho = \rho_1 + \rho_2$ eine Zerlegung gemäß Lemma 3.16 .

$\rho = \mathcal{A}_1^{-1}\rho_1 + \mathcal{A}_1^{-1}\rho_2 = \mathcal{A}_1^{-1}A_1\mathcal{A}_1\rho + \mathcal{A}_1^{-1}\rho_2$, wobei A_1 die orthogonale Projektion auf $ca(im\rho_1)$ ist. Mithin ist
$\Lambda\rho = \Lambda\mathcal{A}_1^{-1}A_1\mathcal{A}_1\rho + \Lambda\mathcal{A}_1^{-1}\rho_2$, also

$$nn_1(w) \leq rg\,\Lambda\mathcal{A}_1^{-1}A_1\mathcal{A}_1 \leq nn_1(\rho).$$

Andererseits sei $\mathcal{A}_2 w = w_1 + w_2$ und folglich $\Lambda\rho = w$
$= \mathcal{A}_2^{-1}w_1 + \mathcal{A}_2^{-1}w_2 = \mathcal{A}_2^{-1}A_2\mathcal{A}_2 w + \mathcal{A}_2^{-1}w_2 = \mathcal{A}_2^{-1}A_2\mathcal{A}_2\Lambda\rho + \mathcal{A}_2^{-1}w_2$,
wobei A_2 auf $ca(imw_1)$ projiziert. Da die $n\times(n-m)$-Matrix Λ den Rang $n-m$ hat, läßt sich eine $(n-m)\times n$-Matrix P finden, so daß $P\Lambda = I_{n-m}$. Damit resultiert
$\rho = P\mathcal{A}_2^{-1}A_2\mathcal{A}_2\Lambda\rho + P\mathcal{A}_2^{-1}w_2$ und deshalb
$nn_1(\rho) \leq rg\,P\mathcal{A}_2^{-1}A_2\mathcal{A}_2\Lambda \leq rg\,\mathcal{A}_2^{-1}A_2\mathcal{A}_2 \leq nn_1(w)$. ▲

Nach diesen Vorbereitungen wenden wir uns der Verallgemeinerung des vollständigen Reiersøl-Resultats zu (siehe S.9 oben), d.h. der Angabe einer notwendigen und hinreichenden Bedingung zur Identifizierbarkeit von Π im multivariaten Fall.

Jeeves[1954] gibt zwar ein solches Resultat an, doch ist sein Beweis an entscheidender Stelle fehlerhaft§, und es ist mir nicht gelungen, seine Beweisidee zu retten. Wir wählen deshalb einen anderen Weg, müssen uns aber auf einen etwas spezielleren Fall beschränken.

---

§ Jeeves Ergebnisse beruhen entscheidend auf seinem Lemma 2: "If A is a symmetric matrix, E is a diagonal matrix with ones and zeros on the main diagonal, and EAE is positive semidefinite, then there exist matrices C, G, and H such that CAC' = DGD + EHE, where D + E = I(the identity) and H is a diagonal matrix with ones and zeros on the main diagonal, C is nonsingular and CD = D." (Da Jeeves Vektoren als Zeilenvektoren schreibt, müssen die Aussagen für unsere Schreibweise geändert werden in C'AC = DGD + EHE, wobei DC = D usw.)
Dieses Ergebnis ist aber falsch, da die geforderte Zerlegung durch passende Wahl von C nur möglich ist, wenn der Block von A, der der Einheitsmatrix in E entspricht, nichtsingulär ist. Dies ist aber nicht der Fall in dem Zusammenhang, in dem Lemma 2 benutzt wird.

Um die Anschauung nicht zu verlieren, nehmen wir vorerst an, daß $\mathrm{im}z \subset \mathbb{R}^3$ und $\mathrm{im}\xi \subset \mathbb{R}^2$. Wir betrachten Modell $\mathfrak{M}_1$ also für n = 3 und m = 1 und kennzeichnen dies durch die Schreibweise $\mathfrak{M}_1(3,1)$.

Satz 3.19

In $\mathfrak{M}_1(3,1)$ mögen alle auftretenden Zufallsvariablen Dichten besitzen. Dann gilt:

$$nn_1(z) = 2 \quad \Rightarrow \quad \langle \Lambda \rangle \text{ ist identifizierbar i.s.S.}$$

Beweis: Wir nehmen an, daß $nn_1(z) = 2$ und $\langle \Lambda \rangle$ nicht identifizierbar wäre, d.h. es gäbe zwei beobachtungsäquivalente Strukturen $\sigma = (\Lambda, F_\xi, F_\nu)$ und $\tilde{\sigma} = (\tilde{\Lambda}, \tilde{F}_\xi, \tilde{F}_\nu)$ mit $\langle \tilde{\Lambda} \rangle \neq \langle \Lambda \rangle$.

Wegen Lemma 3.17 gilt für die Struktur σ:

$$2 = nn_1(z) = nn_1(\Lambda\xi + \nu) \leq nn_1(\Lambda\xi) \leq \mathrm{rg}\,\Lambda = 2 ,$$

also $\quad nn_1(\zeta) = nn_1(\Lambda\xi) = \dim\langle\Lambda\rangle$

und in entsprechender Weise für die Struktur $\tilde{\sigma}$:

$$nn_1(\tilde{\zeta}) = nn_1(\tilde{\Lambda}\xi) = \dim\langle\tilde{\Lambda}\rangle .$$

z hätte also zwei Darstellungen

$$z = u + v$$

und

$$z = \tilde{u} + \tilde{v} ,$$

wobei einmal $\mathrm{ca}(\mathrm{im}u) = \mathrm{ca}(\mathrm{im}\zeta) = \langle\Lambda\rangle$ und $v = \nu$ und das andere Mal $\mathrm{ca}(\mathrm{im}\tilde{u}) = \mathrm{ca}(\mathrm{im}\tilde{\zeta}) = \langle\tilde{\Lambda}\rangle$ und $\tilde{v} = \tilde{\nu}$.

Wir nutzen Lemma 3.16 und wählen eine Scherung $\mathcal{A}_2$ so, daß

$$\mathcal{A}_2 z = \tilde{z}_1 + \tilde{z}_2 , \tag{3.7}$$

$\mathrm{ca}(\mathrm{im}\tilde{z}_1) = \langle\tilde{\Lambda}\rangle$, $\tilde{z}_1$ und $\tilde{z}_2$ stochastisch unabhängig sind, $\tilde{z}_2$ normal ist und $\mathrm{ca}(\mathrm{im}\tilde{z}_2) \perp \langle\tilde{\Lambda}\rangle$ und eine weitere Scherung $\mathcal{A}_1$ so, daß

$$\mathcal{A}_1 \mathcal{A}_2 z = \mathcal{A}_1(\mathcal{A}_2 u + \mathcal{A}_2 v) = z_1 + z_2 \tag{3.8}$$

mit $\mathrm{ca}(\mathrm{im}z_1) = \langle\mathcal{A}_2\Lambda\rangle$, z_1 und z_2 stochastisch unabhängig, z_2 normal und $\mathrm{ca}(\mathrm{im}z_2) \perp \langle\mathcal{A}_2\Lambda\rangle$. Da $\langle\Lambda\rangle \neq \langle\tilde{\Lambda}\rangle$ und $\mathcal{A}_1$ und $\mathcal{A}_2$ bijektiv sind, ist auch $\langle\mathcal{A}_2\Lambda\rangle \neq \langle\mathcal{A}_1\tilde{\Lambda}\rangle$.

Wir unterscheiden zwei Fälle, je nachdem ob

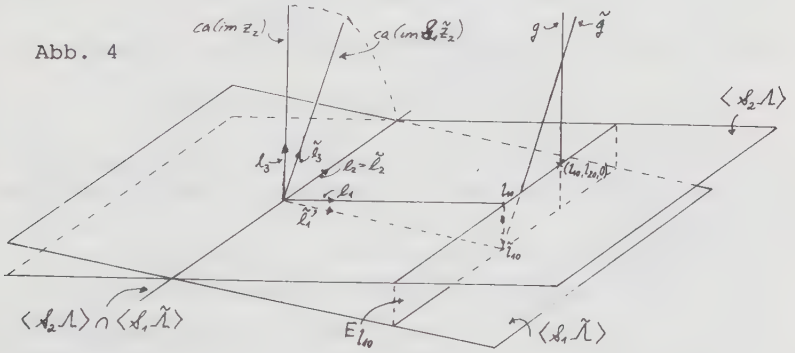
$$\langle\mathcal{A}_2\Lambda\rangle \cap \langle\mathcal{A}_1\tilde{\Lambda}\rangle \subset \langle \mathrm{ca}(\mathrm{im}z_2),\ \mathrm{ca}(\mathrm{im}\,\mathcal{A}_1\tilde{z}_2)\rangle$$

ist (Fall 1) oder nicht (Fall 2).

Die Idee des Beweises besteht darin, die durch die unterschiedlichen Strukturen erzeugten Dichten zu vergleichen und dabei Widersprüche herzuleiten. Zur Darstellung der Dichten ist es vorteilhaft, den Strukturen und Fällen entsprechend im $\mathbf{R}^3$ verschiedene Basen zu wählen.

Wir behandeln nun zunächst den Fall 1. Da $\langle \mathcal{A}_2 \Lambda \rangle \cap \langle \mathcal{A}_1 \Lambda \rangle \neq \{0\}$ muß dann auch $ca(im z_2) \neq \{0\}$ und $ca(im \mathcal{A}_1 \tilde{z}_2) \neq \{0\}$ sein. Außerdem kann weder $ca(im z_2)$ noch $ca(im \mathcal{A}_1 \tilde{z}_2)$ in $\langle \mathcal{A}_2 \Lambda \rangle \cap \langle \mathcal{A}_1 \Lambda \rangle$ liegen. Wir wählen für diesen Fall zwei Sätze ℓ_1, ℓ_2, ℓ_3 und $\mathfrak{l}_1, \mathfrak{l}_2, \mathfrak{l}_3$ von Basisvektoren auf folgende Weise. Der Vektor ℓ_3 sei kollinear mit $ca(im z_2)$, $\mathfrak{l}_3$ kollinear mit $ca(im \mathcal{A}_1 \tilde{z}_2)$. Weiterhin seien $\ell_2 = \mathfrak{l}_2$ kollinear mit $\langle \mathcal{A}_2 \Lambda \rangle \cap \langle \mathcal{A}_1 \Lambda \rangle$. Schließlich sollen ℓ_1 bzw. $\mathfrak{l}_1$ in $\langle \mathcal{A}_2 \Lambda \rangle$ bzw. $\langle \mathcal{A}_1 \Lambda \rangle$ liegen und orthogonal zu ℓ_2 sein. (siehe Abb. 4)

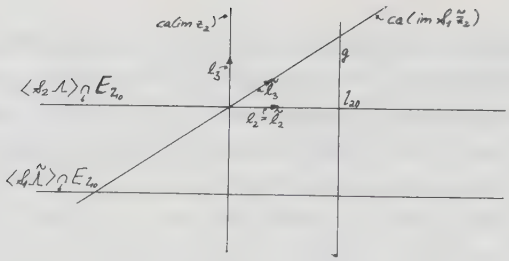
Abb. 4



Die Basen $(\ell_1, \ell_2, \ell_3), (\mathfrak{l}_1, \mathfrak{l}_2, \mathfrak{l}_3)$ lassen sich gemäß $(\mathfrak{l}_1, \mathfrak{l}_2, \mathfrak{l}_3) = (\ell_1, \ell_2, \ell_3)T'$ mit $T' = \begin{pmatrix} t_{11} & 0 & 0 \\ 0 & 1 & t_{32} \\ t_{13} & 0 & t_{33} \end{pmatrix}$ transformieren.

Die entsprechende Koordinatentransformation ergibt sich daraus zu $(I_1, I_2, I_3) = (l_1, l_2, l_3)T^{-1}$, wobei T^{-1} die Gestalt $\begin{pmatrix} t^{11} & t^{12} & t^{13} \\ 0 & 1 & 0 \\ 0 & t^{32} & t^{33} \end{pmatrix}$ besitzt.

In Abbildung 4 ist die Ebene $E_{l_{10}}$ angedeutet. Als Aufsicht auf diese Ebene erhalten wir folgendes Bild:



Es sei l_{10} so gewählt, daß die nachfolgenden Dichten Masse haben. Ein solches l_{10} existiert immer, da sonst nicht $nn_1(z) = 2$ sein könnte.

Die Dichte von z läßt sich nun auf zwei Weisen erzeugen. Für σ und die Basis (ℓ_1,ℓ_2,ℓ_3) ergibt sich:

$$f_{\mathcal{L}_1\mathcal{L}_2 Z}(l_1,l_2,l_3) = f_{Z_1}(l_1,l_2,0)\cdot f_{Z_2}(0,0,l_3) \ .$$

Für $\tilde{\sigma}$ und die Basis $(\tilde{\ell}_1,\tilde{\ell}_2,\tilde{\ell}_3)$ erhalten wir:

$$f_{\tilde{\mathcal{L}}_1\tilde{\mathcal{L}}_2 Z}(\tilde{l}_1,\tilde{l}_2,\tilde{l}_3)=f_{\mathcal{L}_1\mathcal{L}_2 Z}(\tilde{l}_1(l_1),\tilde{l}_2(l_1,l_2,l_3),\tilde{l}_3(l_1,l_3))$$

$$=f_{\mathcal{L}_1\tilde{Z}_1}(\tilde{l}_1,\tilde{l}_2,0)\cdot f_{\mathcal{L}_1\tilde{Z}_2}(0,0,\tilde{l}_3) \ .$$

Betrachten wir speziell die Dichte auf der Geraden

$$g=\{(l_{10},l_{20},l_3)\,|\,l_3\in\mathbb{R}\}=\{(\tilde{l}_1(l_{10}),\tilde{l}_2(l_{10},l_{20},l_3),\tilde{l}_3(l_{10},l_3))\,|\,l_3\in\mathbb{R}\}$$

so resultiert

$$f_{Z_1}(l_{10},l_{20},0)\cdot f_{Z_2}(0,0,l_3)=f_{\mathcal{L}_1\tilde{Z}_1}(\tilde{l}_1(l_{10}),\tilde{l}_2(l_{10},l_{20},l_3),0)\ \times$$

$$\times\ f_{\mathcal{L}_1\tilde{Z}_2}(0,0,\tilde{l}_3(l_{10},l_3))$$

bzw.

$$f_{\mathcal{L}_1\tilde{Z}_1}(\tilde{l}_1(l_{10}),\tilde{l}_2(l_{10},l_{20},l_3),0)=f_{Z_1}(l_{10},l_{20},0)\,f_{Z_2}(0,0,l_3)\ \times$$

$$\times\ f^{-1}_{\mathcal{L}_1\tilde{Z}_2}(0,0,\tilde{l}_3(l_{10},l_3)) \ . \qquad (3.9)$$

Die Menge $\{(\tilde{l}_1(l_{10}),\tilde{l}_2(l_{10},l_{20},l_3),0)\,|\,l_3\in\mathbb{R}\}$ ist die Gerade $\langle\mathcal{L}_1\tilde{\mathcal{L}}\rangle \cap E_{l_{10}}$.

$f_{Z_1}(0,0,l_3)$ und $f_{\mathcal{L}_1\tilde{Z}_2}(0,0,\tilde{l}_3(l_{10},l_3))$ sind Dichten von Normalverteilungen (kurz: normale Dichten), so daß die Dichte (3.9) auf $\langle\mathcal{L}_1\tilde{\mathcal{L}}\rangle \cap E_{l_{10}}$ ebenfalls eine normale Dichte ist. Diese Folgerung , die wir in ähnlicher Weise im weiteren Verlauf des Beweises noch öfter machen werden, ist

durch folgenden Satz von Marcinkiewicz(siehe z.B. Kagan, Linnik, Rao[1973], Lemma 1.4.2) gerechtfertigt:

"Hat die c.F. $\phi(t)$ die Form $\exp[P(t)]$, wobei P ein Polynom ist, so ist P ein quadratisches Polynom und ϕ die c.F. einer Normalverteilung."

Betrachten wir andererseits speziell die Dichte auf $\tilde{g} = \{(\tilde{l}_{10}, 0, \tilde{l}_3) \mid \tilde{l}_3 \in \mathbb{R}\}$, so bekommen wir aus

$$f_{z_1}(l_1(\tilde{l}_{10}), l_2(\tilde{l}_3), 0) f_{z_2}(0,0,l_3(\tilde{l}_{10}, \tilde{l}_3)) = f_{\not{A}_1 \tilde{z}_1}(\tilde{l}_{10}, 0, 0) f_{\not{A}_1 \tilde{z}_2}(0,0,\tilde{l}_3)$$

als Dichte auf der Geraden $\langle \not{A}_2 \Lambda \rangle \cap E_{l_{10}}$

$$f_{z_1}(l_1(\tilde{l}_{10}), l_2(\tilde{l}_3), 0) = f_{\not{A}_1 \tilde{z}_1}(\tilde{l}_{10}, 0, 0) f_{\not{A}_1 \tilde{z}_2}(0,0,\tilde{l}_3) f^{-1}_{z_2}(0,0,l_3(\tilde{l}_{10}, \tilde{l}_3)).$$

Diese ist wiederum normal.

$f_{\not{A}_1 \tilde{z}_1}(\tilde{l}_1(l_{10}), \tilde{l}_2(l_{10}, l_{20}, l_3), 0)$ und $f_{\not{A}_1 \tilde{z}_2}(0,0,\tilde{l}_3)$ bzw. $f_{z_1}(l_1(\tilde{l}_{10}), l_2(\tilde{l}_3), 0)$ und $f_{z_2}(0,0,l_3)$ erzeugen normale Dichten auf $E_{l_{10}}$. Diese stehen aber im Widerspruch zueinander. Denn erstens werden diese Verteilungen wegen der Normalität durch ihre Konzentrationsellipsoide vollständig charakterisiert. Zweitens müssen aber wegen der Unabhängigkeit von z_1 und z_2 bzw. $\not{A}_1 \tilde{z}_1$ und $\not{A}_1 \tilde{z}_2$ die Richtungen der Träger der unabhängigen Komponenten konjugiert bezüglich des Konzentrationsellipsoids sein. Dies kann aber nicht der Fall sein, da $\langle \not{A}_2 \Lambda \rangle \cap E_{l_{10}}$ kollinear zu $\langle \not{A}_1 \Lambda \rangle \cap E_{l_{10}}$ ist, aber $ca(im z_2)$ nicht zu $ca(im \not{A}_1 \tilde{z}_2)$. Somit ist Fall 1 ausgeschlossen.

Wenden wir uns dem Fall 2 zu. Da $nn_1(z) = 2$, $\dim \langle \Lambda \rangle = 2$ und σ und $\tilde{\sigma}$ beobachtungsäquivalent, aber verschieden sind, können wir ausschließen, daß $ca(im z_2) = 0$ v $ca(im \not{A}_1 \tilde{z}_2) = 0$.

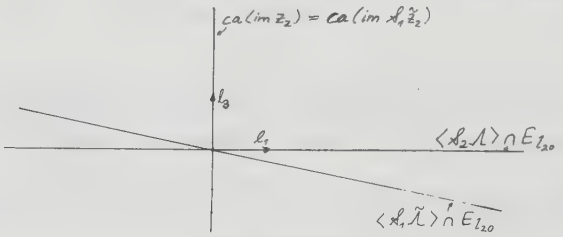
Wir zeigen nun als erstes, daß

$$ca(im z_2) = ca(im \not{A}_1 \tilde{z}_2)$$

ausgeschlossen ist.

Angenommen $ca(imz_2)$ und $ca(im\tilde{z}_2)$ wären gleich. Wir behalten ℓ_1, ℓ_2, ℓ_3 als Basis und betrachten den Schnitt bei $l_2 = l_{20}$.

Abb. 6



Dadurch können wir uns auf Betrachtungen im $\mathbb{R}^2$ beschränken. Dieser sei nun durch ℓ_1 und ℓ_3 aufgespannt. Die Zerlegungen (3.7) und (3.8) werden in diesem Fall zu

$$z_1(\cdot)\begin{pmatrix}1\\0\end{pmatrix} + z_2(\cdot)\begin{pmatrix}0\\1\end{pmatrix} = \tilde{z}_1(\cdot)\begin{pmatrix}1\\\lambda\end{pmatrix} + \tilde{z}_2(\cdot)\begin{pmatrix}0\\1\end{pmatrix} , \quad (3.10)$$

wobei hier in etwas mißbräuchlicher Schreibweise $z_1(\cdot)$, $z_2(\cdot)$, $\tilde{z}_1(\cdot)$ und $\tilde{z}_2(\cdot)$ reellwertige Zufallsvariablen bezeichnen, die zu z_1, z_2, $\tilde{z}_1$ und $\tilde{z}_2$ aus (3.7) und (3.8) 'passen' und $\lambda \neq 0$ ist.

In c.F. ausgedrückt, lautet (3.10)

$$\phi_{z_1(\cdot)}(t_1) \cdot \phi_{z_2(\cdot)}(t_2) = \phi_{\tilde{z}_1(\cdot)}(t_1 - \lambda t_2) \cdot \phi_{\tilde{z}_2(\cdot)}(t_2) .$$

Wir projizieren in zwei Weisen auf $ca(imz_2) = ca(im\tilde{z}_2)$; einmal entlang $\langle \ell_2 \Lambda \rangle \cap E_{l_{20}}$ (d.h. Projektionsmatrix $P_1 = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$), das andere Mal entlang $\langle \ell_1 \tilde{\Lambda} \rangle \cap E_{l_{20}}$ ($P_2 = \begin{pmatrix}0 & 0\\\lambda & 1\end{pmatrix}$) und erhalten nach P_1:

$$\phi_{z_1(\cdot)}(0) \cdot \phi_{z_2(\cdot)}(t_2)$$

$$= \phi_{\tilde{z}_1(\cdot)}(-\lambda t_2) \cdot \phi_{\tilde{z}_2(\cdot)}(t_2)$$

d.h.

$$\phi_{\tilde{z}_1(\cdot)}(-\lambda t_2) = \phi_{z_2(\cdot)}(t_2) \cdot \phi^{-1}_{\tilde{z}_2(\cdot)}(t_2) \quad (3.11)$$

bzw. nach P_2:

$$\phi_{z_1(\cdot)}(\lambda\, t_2)\cdot\phi_{z_2(\cdot)}(t_2)$$

$$= \phi_{\tilde{z}_1(\cdot)}(0)\cdot\phi_{\tilde{z}_2(\cdot)}(t_2)$$

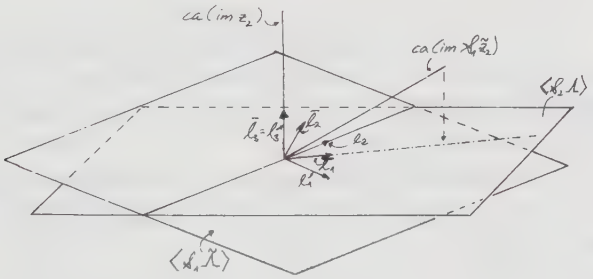
d.h.

$$\phi_{z_1(\cdot)}(\lambda\, t_2) = \phi_{\tilde{z}_2(\cdot)}(t_2)\cdot\phi^{-1}_{z_2(\cdot)}(t_2) \qquad (3.12)$$

(3.11) und (3.12) führen zu einem Widerspruch zu $nn_1(z)=2$.

Somit muß $ca(imz_2) \neq ca(im\chi_1 \tilde{z}_2)$ sein. Wir führen nun eine weitere Basis $\mathfrak{l}_1, \mathfrak{l}_2, \mathfrak{l}_3$ ein. $\mathfrak{l}_1$ sei kollinear mit der Projektion von $ca(im\chi_1 \tilde{z}_2)$ entlang $ca(imz_2)$ auf $\langle \chi_2 \Lambda \rangle$, $\mathfrak{l}_3$ sei kollinear mit $ca(imz_2)$ und $\mathfrak{l}_2$ sei orthogonal zu $\mathfrak{l}_1$ und $\mathfrak{l}_3$ (siehe Abb. 7).

Abb. 7



Durch eine dritte Scherung $\chi_3 : \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} \to \begin{pmatrix} I_1 - kI_2 \\ I_2 \\ I_3 \end{pmatrix}$,

wobei k passend zu wählen ist, erreichen wir, daß $\mathfrak{l}_2$ mit $\langle \chi_3 \chi_1 \Lambda \rangle \cap \langle \chi_3 \chi_2 \Lambda \rangle$ kollinear ist.§

Wir unterscheiden, ob

$$ca(imz_2) \subset \langle \chi_1 \Lambda \rangle \;\vee\; ca(im\chi_1 \tilde{z}_2) \subset \langle \chi_2 \Lambda \rangle \qquad (3.13)$$

gilt (Fall 2a) oder nicht (Fall 2b).

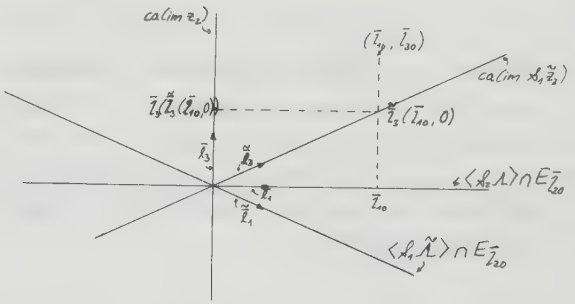
Gilt nur eine der Möglichkeiten von (3.13), so zeigt man wie oben, daß daraus ein Widerspruch erwächst. Gelten beide Möglichkeiten, so ist klar, daß in $E_{\mathfrak{l}_2 0}$ die bedingte Dichte normal ist. Da $ca(imz_2)$ und $ca(im\chi_1 \tilde{z}_2)$ unabhängig von I_2 sind, ist die Randdichte auf der $\mathfrak{l}_1, \mathfrak{l}_3$-Ebene normal.

§ Um die Bezeichnungen im folgenden nicht mehr als nötig zu komplizieren, lassen wir das Zeichen χ_3 wieder fallen.

Dies führt zu einem Widerspruch zur Annahme, daß $nn_1(z)=2$.

Es bleibt der Fall 2b. Für festes $\bar{I}_{20}$ haben wir folgendes Bild:

Abb. 8



Zur Darstellung der durch die Struktur $\tilde{\sigma}$ erzeugten Dichten führen wir weitere Basisvektoren $\tilde{\bar{\iota}}_1$ $\tilde{\bar{\iota}}_2$ $\tilde{\bar{\iota}}_3$ ein, wobei $\tilde{\bar{\iota}}_3$ kollinear mit $ca(im\,\ell_1\,\tilde{z}_2)$, $\tilde{\bar{\iota}}_2 = \bar{\iota}_2$ und $\tilde{\bar{\iota}}_1$ kollinear mit $\langle \ell_1 \tilde{\ell} \rangle \cap E_{\bar{I}_{20}}$ sind.

Wir betrachten die Dichten in Punkten der Menge $\{(0,\tilde{\bar{I}}_{20},\tilde{\bar{I}}_3) \mid \tilde{\bar{I}}_3 \in \mathbb{R}\} = \{(\bar{I}_1(0,\tilde{\bar{I}}_3),\bar{I}_2,\bar{I}_3(0,\tilde{\bar{I}}_3)) \mid \tilde{\bar{I}}_3 \in \mathbb{R}\}$ und erhalten:

$$f_{z_1}(\bar{I}_1(\tilde{\bar{I}}_3),\bar{I}_{20},0)\, f_{z_2}(0,0,\bar{I}_3(\tilde{\bar{I}}_3)) = f_{\ell_1 \tilde{z}_1}(0,\tilde{\bar{I}}_{20},0)\, f_{\ell_1 \tilde{z}_2}(0,0,\tilde{\bar{I}}_3)$$

bzw.

$$f_{z_1}(\bar{I}_1(\tilde{\bar{I}}_3),\bar{I}_{20},0) = f_{\ell_1 \tilde{z}_1}(0,\tilde{\bar{I}}_{20},0)\, f_{\ell_1 \tilde{z}_2}(0,0,\tilde{\bar{I}}_3)\, f_{z_2}^{-1}(0,0,\bar{I}_3(\tilde{\bar{I}}_3)) \quad .$$

Die bedingte Dichte auf der $\bar{\iota}_1,\bar{\iota}_3$-Ebene für gegebenes $\bar{I}_{20}$, ergibt sich zu

$$= \frac{f_{z_1}(\bar{I}_{10},\bar{I}_{20},0)\, f_{z_2}(0,0,\bar{I}_{30})}{\int\int(\ldots.)\,d\bar{I}_{10}\,d\bar{I}_{30}}$$

$$= \frac{f_{\ell_1 \tilde{z}_1}(0,\tilde{\bar{I}}_{20},0)\, f_{z_2}(0,0,\tilde{\bar{I}}_3(\bar{I}_{10},0))\, f_{z_2}^{-1}(0,0,\bar{I}_3(\tilde{\bar{I}}_3(\bar{I}_{10},0)))\, f_{z_2}(0,0,\bar{I}_{30})}{\int\int(\ldots.)\,d\bar{I}_{10}\,d\bar{I}_{30}}$$

$$= \frac{f_{z_2}(0,0,\tilde{\bar{I}}_3(\bar{I}_{10},0))\, f_{z_2}^{-1}(0,0,\bar{I}_3(\tilde{\bar{I}}_3(\bar{I}_{10},0)))\, f_{z_2}(0,0,\bar{I}_{30})}{\int\int(\ldots.)\,d\bar{I}_{10}\,d\bar{I}_{30}} \quad ,$$

wobei in die Klammern der Nenner jeweils die Zähler zu setzen sind. Diese Dichte ist unabhängig von $\bar{I}_2$ und normal. Dies führt aber wie vorne zu einem Widerspruch zu $nn_1(z) = 2$. ▲

Wegen der Anschaulichkeit hatten wir uns zunächst auf $imz \subset \mathbb{R}^3$ und $im\xi \subset \mathbb{R}^2$ beschränkt. Der obige Beweis läßt sich aber unmittelbar auf $\mathfrak{M}_1(n,1)$, also $imz \subset \mathbb{R}^n$ und $im\xi \subset \mathbb{R}^{n-1}$ ausdehnen, d.h. in

$$z = \begin{pmatrix} z_1 \\ \cdot \\ \cdot \\ z_{n-1} \\ z_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_{n-1} \\ y \end{pmatrix} = \begin{pmatrix} \xi_1 + \delta_1 \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \xi_{n-1} + \delta_{n-1} \\ \pi_1\xi_1 + \cdots + \pi_{n-1}\xi_{n-1} + \varepsilon \end{pmatrix}$$

sind die Koeffizienten $\pi_1, \ldots, \pi_{n-1}$ i.s.S. identifizierbar, wenn $nn_1(z) = n-1$.

Betrachten wir nun (1.3), wo $imz \subset \mathbb{R}^n$ und $im\xi \subset \mathbb{R}^{n-m}$, und bezeichnen wir mit $z^{(j)}$ den Vektor

$$(z_1, \ldots, z_{n-m}, z_{n-m+j})' = (x_1, \ldots, x_{n-m}, y_j)',$$

so implizieren die obigen Überlegungen

Satz 3.20

In $\mathfrak{M}_1$ gilt unter der Annahme des Satzes 3.5:

$$\bigwedge_{j=1,\ldots,m} nn_1(z^{(j)}) = n-m \Rightarrow \langle\Lambda\rangle \text{ i.s.S. identifizierbar.}$$

Auf der anderen Seite erhebt sich die Frage, ob zunächst in Satz 3.19 auch die umgekehrte Implikation gilt, d.h. ob $nn_1(z)=2$ auch notwendig zur Identifizierbarkeit von $\langle\Lambda\rangle$ ist.

Wir können dies in folgender Weise bejahen. Ist $nn_1(z)<2$, so existiert eine Darstellung $z = u + v$ mit $\dim ca(imu)<2$ und v unabhängig von u normalverteilt. Ist $nn_1(z)=0$, so ist aus dem Beispiel in Abschnitt 1.2 klar, daß $\langle\Lambda\rangle$ nicht identifizierbar ist. Also sei $nn_1(z)=1$.

Wir nehmen an, daß keine Variable exakt beobachtbar ist und, daß dim ca(imz) = 3 ist. Der normalverteilte Vektor v sei so in zwei Komponenten $v_1 + v_2$ zerlegt, daß $ca(imv_1) = ca(imu)$ und v_1, v_2 stochastisch unabhängig sind. Der eindimensionale Träger von u werde durch den Vektor 1 erzeugt.

Wir können nun beliebig viele Ebenen wählen, die durch 1 und solche Vektoren k aus $ca(imv_2)$ aufgespannt werden, die erlauben, daß die Ebenen auch durch Vektoren der Form $(1,0,\tilde{\pi}_1)'$ und $(0,1,\tilde{\pi}_2)'$ erzeugt werden können. Mit $v_2 = v_{21}(\cdot)k + v_{22}(\cdot)\bar{k}$, wobei $v_{21}(\cdot)$ und $v_{22}(\cdot)$ unabhängige reellwertige Zufallsvariablen sind, gilt:

$$
\begin{aligned}
z = u + v &= u(\cdot)1 + v_1(\cdot)1 + v_{21}(\cdot)k + v_{22}(\cdot)\bar{k} \\
&= (1,k)\begin{pmatrix} u(\cdot) + v_1(\cdot) \\ v_{21}(\cdot) \end{pmatrix} + v_{22}(\cdot)\bar{k} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \tilde{\pi}_1 & \tilde{\pi}_2 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \tilde{v} \quad ,
\end{aligned}
$$

ξ_1 und ξ_2 sind Linearkombinationen von $u(\cdot), v_1(\cdot)$ und $v_{21}(\cdot)$ und somit unabhängig von $\tilde{v} = v_{22}(\cdot)k$. Die entsprechenden Strukturen sind somit zulässig und beobachtungsäquivalent.

Das obige läßt sich sofort auf höhere Dimensionen ausdehnen, d.h. es gilt

Satz 3.21

Gegeben sei $\mathfrak{M}_1$. Ferner sei keine Variable exakt beobachtbar und dim ca(imz) = n. Dann gilt:

$$\Lambda \text{ i.s.S. identifizierbar} \Rightarrow nn_1(z) = n-m \ .$$

Es bleibt die Frage, ob die hinreichende Bedingung aus Satz 3.20 äquivalent ist zur notwendigen Bedingung aus Satz 3.21, d.h. ob aus $nn_1(z) = n-m$ folgt, daß $nn_1(z^{(j)}) = n-m$ für alle $j=1,\ldots,m$ ist.

Hierzu wissen wir:

$$nn_1(z) = n-m \;\Rightarrow\; nn_1(\xi) = n-m$$

$$\Rightarrow\; nn_1(z^{(j)}) = nn_1\left(\begin{pmatrix} I \\ \Pi_{j.} \end{pmatrix}\xi+\nu^{(j)}\right)$$

$$\underset{(*)}{\leq} nn_1\left(\begin{pmatrix} I \\ \Pi_{j.} \end{pmatrix}\xi\right) = nn_1(\xi) = n-m \quad .$$

Die Ungleichheit (*) folgt aus Lemma 3.17. Würde hier Gleichheit gelten, wären die Bedingungen äquivalent. Ob dies der Fall ist, bleibt, wie bereits vorne gesagt, ein offenes Problem.

Zum Abschluß dieses Unterabschnitts wollen wir eine weitere hinreichende Bedingung zur Identifizierbarkeit von Π vorstellen, die mit Hilfe einer alternativen Definition der Nichtnormalität formuliert wird und aus einem 'konstruktiven' Identifikationsansatz von Jeeves resultiert.

Die Idee beruht auf folgender Beobachtung in $\mathfrak{M}_1$:

Lemma 3.22

In $\mathfrak{M}_1$ gilt für jeden Projektor P nach $\langle\Lambda\rangle$:

$$\Psi_z(t) = \Psi_z(P't) + f(t) \;,$$

wobei $\Psi(t) = \log\phi(t)$ und $f(t)$ ein quadratisches Polynom ist.

Beweis: In $\mathfrak{M}_1$ gilt: $\Psi_z(t) = \Psi_\xi(\Lambda't) + \Psi_\nu(t)$. Da $\langle P\rangle = \langle\Lambda\rangle$ nach Voraussetzung und somit $\Lambda = PQ$ für eine nichtsinguläre Matrix Q, erhalten wir durch Setzen von P't für t:

$$\Psi_z(P't)=\Psi_\xi(Q'P'P't)+\Psi_\nu(P't)=\Psi_\xi(Q'P't)+\Psi_\nu(P't)=\Psi_\xi(\Lambda't)+\Psi_\nu(P't)$$

und daraus:

$$\Psi_z(t)=\Psi_z(P't)+\Psi_\nu(t)-\Psi_\nu(P't)=\Psi_z(P't)+f(t) \quad . \quad \blacksquare$$

Die Idee ist nun, die Differenz zwischen den kumulantenerzeugenden Funktionen Ψ von z und dessen Projek-

tionen in verschiedene Unterräume zu betrachten und den wahren Unterraum $\langle\Lambda\rangle$ dann dadurch zu identifizieren, daß er der kleinste ist, bei dem die Differenz ein quadratisches Polynom ist. Damit man mit diesem Verfahren auch $\langle\Lambda\rangle$ findet, muß die Nichtnormalität von z entsprechend definiert werden.

Definition 3.23(Jeeves)

Die N i c h t n o r m a l i t ä t i m z w e i t e n S i n n e $nn_2(z)$ von z ist der minimale Rang aller Matrizen A mit $\Psi_z(t) = \Psi_z(A't) + f(t)$.

Wir erhalten damit den

Satz 3.24

In $\mathfrak{M}_1$ gilt: $nn_2(z) = n-m \Rightarrow \langle\Lambda\rangle$ i.s.S. identifizierbar.

Beweis: Angenommen, $\langle\Lambda\rangle$ sei nicht identifizierbar; dann ist zu zeigen, daß $nn_2(z) \neq n-m$ ist, d.h. da $nn_2(z) \leq n-m$ immer gilt, daß $nn_2(z) < n-m$ sein muß. Nehmen wir also an, daß es zwei beobachtungsäquivalente, aber verschiedene Unterräume $\langle\Lambda\rangle$ und $\langle\bar{\Lambda}\rangle$ gibt, d.h $\dim(\langle\Lambda\rangle \cap \langle\bar{\Lambda}\rangle) < n-m$. Nach Lemma 3.22existieren Darstellungen

$$\Psi_z(t) = \Psi_z(P't) + f_1(t) \qquad (3.14)$$

$$\Psi_z(t) = \Psi_z(\bar{P}'t) + f_2(t) \ , \qquad (3.15)$$

wobei P und $\bar{P}$ Projektionen in $\langle\Lambda\rangle$ und $\langle\bar{\Lambda}\rangle$ sind und P entlang $\langle\bar{\Lambda}\rangle$ projizieren soll. Durch Einsetzen von P't für t in (3.15) erhalten wir zunächst

$$\Psi_z(P't) = \Psi_z(\bar{P}'P't) + f_2(P't)$$

und in Verbindung mit (3.14)

$$\Psi_z(t) = \Psi_z(\bar{P}'P't) + f_3(t) \ .$$

Aber $rg(P\bar{P}) = \dim(\langle\Lambda\rangle \cap \langle\bar{\Lambda}\rangle) < n-m$, womit in Zusammenhang mit Definition 3.23der Satz bewiesen ist. ▲

Bemerkung 3.25

Es gilt $nn_2(z) \leq nn_1(z)$. Denn aus Lemma 3.16 folgt $z = \sqrt{}^{-1} z_1 + \sqrt{}^{1} z_2 = \sqrt{}^{-1} A_1 \sqrt{} z + \sqrt{}^{1} z_2$, bzw.

$$\Psi_z(t) = \Psi_z((\sqrt{}^{-1} A_1 \sqrt{})' t) + f(t) .$$

Nun ist aber $rg(\sqrt{}^{-1} A_1 \sqrt{}) \leq nn_1(z)$, so daß

$$nn_2(z) \leq rg(\sqrt{}^{-1} A_1 \sqrt{}) \leq nn_1(z) .$$

Ob sogar $nn_1(z) = nn_2(z)$ gilt, ist (entgegen der Behauptung von Jeeves) nicht erwiesen.

3.1.4 Identifizierbarkeit des Rests der Struktur

Wir wollen nun untersuchen, ob in $\mathfrak{M}_1$ bei Identifizierbarkeit von Π, die restlichen Elemente der Struktur $\sigma=(\Lambda,\Omega,F_\xi)$ bzw. $(\Lambda,\Omega,\phi_\xi)$ identifizierbar sind. Dazu betrachten wir wieder zwei beobachtungsäquivalente Strukturen, in denen aber jetzt Π fest ist.

Die Kovarianzmatrix Ω sei entsprechend der Aufspaltung von ν in δ und ε in Blöcke unterteilt, also

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} .$$

$$\phi_\xi(t_1+\Pi' t_2)\cdot\exp\{-1/2(t_1'\Omega_{11}t_1+t_1'\Omega_{12}t_2+t_2'\Omega_{21}t_1+t_2'\Omega_{22}t_2)\}$$

$$= \tilde{\phi}_\xi(t_1+\Pi' t_2)\cdot\exp\{-1/2(t_1'\tilde{\Omega}_{11}t_1+t_1'\tilde{\Omega}_{12}t_2+t_2'\tilde{\Omega}_{21}t_1+t_2'\tilde{\Omega}_{22}t_2)\}$$

bzw. durch Logarithmieren:

$$\Psi_\xi(t_1+\Pi' t_2)-\Psi_\xi(t_1+\Pi' t_2) = -1/2[t_1'(\tilde{\Omega}_{11}-\Omega_{11})t_1+t_1'(\tilde{\Omega}_{12}-\Omega_{12})t_2$$
$$+t_2'(\tilde{\Omega}_{21}-\Omega_{21})t_1+t_2'(\tilde{\Omega}_{22}-\Omega_{22})t_2].$$

Wir setzen $t_1+\Pi' t_2 = s$ und eliminieren t_1:

$$\Psi_\xi(s)-\Psi_\xi(s) = -1/2[(s-\Pi' t_2)'(\Delta\Omega_{11})(s-\Pi' t_2)+(s-\Pi' t_2)(\Delta\Omega_{12})t_2$$
$$+t_2'(\Delta\Omega_{21})(s-\Pi' t_2)+t_2'(\Delta\Omega_{22})t_2] \qquad (3.16)$$

$$= -1/2\{s'[(\Delta\Omega_{11})(-\Pi')+(\Delta\Omega_{12})]t_2+t_2'[-\Pi(\Delta\Omega_{11})+(\Delta\Omega_{21})]s$$

$$+t_2'[\Pi(\Delta\Omega_{11})\Pi'-\Pi(\Delta\Omega_{12})-(\Delta\Omega_{21})\Pi'+(\Delta\Omega_{22})]t_2$$

$$+s'(\Delta\Omega_{11})s\} ,$$

wobei $\Delta\Omega_{ij}$ für $\tilde{\Omega}_{ij} - \Omega_{ij}$ gesetzt wurde.
Da dies eine Identität in s und t_2 ist, muß

$$(\Delta\Omega_{11})\Pi' = \Delta\Omega_{12} \qquad (3.17)$$

gelten und wegen der Symmetrie der Matrix in eckigen Klammern in der obigen zweiten Zeile

$$\Pi(\Delta\Omega_{11})\Pi'-\Pi(\Delta\Omega_{12})-(\Delta\Omega_{21})\Pi'+\Delta\Omega_{22} = 0 \qquad (3.18')$$

sein, bzw. durch Kombination von (3.17) und (3.18') auch

$$(\Delta\Omega_{21})\Pi' = \Delta\Omega_{22} \qquad (3.18)$$

gelten. Schließlich ergeben (3.17) und (3.18) in (3.16) eingesetzt

$$\Psi_\xi(s)-\tilde{\Psi}_\xi(s) = -1/2[s'(\Delta\Omega_{11})s]$$

bzw. $$\tilde{\phi}_\xi(s) = \phi_\xi(s)\cdot\exp\{+1/2[s'(\Delta\Omega_{11})s]\} . \qquad (3.19)$$

Mit (3.17),(3.18) und (3.19) können wir nun von $\Omega_{11},\Omega_{12},\Omega_{22},\phi_\xi$ verschiedene, aber dazu beobachtungsäquivalente $\tilde{\Omega}_{11},\tilde{\Omega}_{12},\tilde{\Omega}_{22},\tilde{\phi}_\xi$ finden. Eine Wahl von $\tilde{\Omega}_{11}$, sodaß $\Delta\Omega_{11}$ positiv definit ist, ist möglich, wenn $\phi_\xi(s)$ einen normalen Faktor besitzt und $\Delta\Omega_{11}$ genügend klein ist. $\tilde{\Omega}_{12},\tilde{\Omega}_{22},\tilde{\phi}_\xi$ ergeben sich dann aus (3.17)-(3.19). Die Wahl eines positiv definiten $\tilde{\Omega}_{11}$ mit $-\Delta\Omega_{11}$ positiv definit ist möglich, wenn Ω_{22} positiv definit ist und $-\Delta\Omega_{11}$ genügend klein ist. Somit ergibt sich die

Bemerkung 3.26

In $\mathfrak{M}_1$ sind selbst bei Identifizierbarkeit von Π Ω und ϕ_ξ i.a. nicht identifizierbar.

Eine Verschärfung der Annahme (C) ändert dies. Mit $\mathfrak{M}_2$ sei das Modell bezeichnet, in dem über die Annahmen in $\mathfrak{M}_1$ hinaus angenommen wird, daß Ω blockdiagonal sei.

Die Blockdiagonalität von Ω hat auch Auswirkungen auf die Identifizierbarkeit von Π, die wir aber nicht weiter untersuchen wollen. Es sei nur erwähnt, daß obgleich sie nicht vollständige Identifizierbarkeit liefert, sie vermutlich den Bereich der Nichtidentifizierbarkeit einschränken wird. Im Fall n=2,m=1 z.B. weiß man, daß das wahre π zwischen den beiden normalen Regressionskoeffizienten liegen muß. Weiteres hierzu entnehme man Frisch [1934](S. 58-65).

Wir nehmen an, daß Π identifizierbar ist. In diesem Fall gilt dann

$$\Psi_\xi(t_1+\Pi' t_2)+\Psi_\delta(t_1)+\Psi_\varepsilon(t_2) = \Phi_\xi(t_1+\Pi' t_2)+\Phi_\delta(t_1)+\Phi_\varepsilon(t_2) \quad (3.20)$$

wobei daran erinnert sei, daß die Ψ's in der Umgebung von Null endlich sind. Wir führen in Anlehnung an Reiersøl den Differenzenoperator

$$\Delta_{\tau_1}\Psi(t_1) = \Psi(t_1+\tau_1) - \Psi(t_1)$$

ein, wobei τ_1 ein (n-m)×1 - Vektor ist. Wir wenden Δ_{τ_1} auf (3.20) an und erhalten

$$\Delta_{\tau_1}\Phi_\xi(t_1+\Pi' t_2)-\Delta_{\tau_1}\Psi_\xi(t_1+\Pi' t_2)+\Delta_{\tau_1}\Phi_\delta(t_1)-\Delta_{\tau_1}\Psi_\delta(t_1) = 0 , \quad (3.21)$$

da $\Delta_{\tau_1}\Psi_\varepsilon(t_2) = 0$. Setzen wir zunächst $t_2=0$ und anschließend $t_1+\Pi' t_2$ anstelle von t_1, so resultiert

$$\Delta_{\tau_1}\Phi_\xi(t_1+\Pi' t_2)-\Delta_{\tau_1}\Psi_\xi(t_1+\Pi' t_2)+\Delta_{\tau_1}\Phi_\delta(t_1+\Pi' t_2)-\Delta_{\tau_1}\Psi_\delta(t_1+\Pi' t_2)=0 \quad (3.22)$$

(3.22) subtrahiert von (3.21) ergibt

$$\Delta_{\tau_1}\Delta_{\Pi' t_2}(\Phi_\delta(t_1)-\Psi_\delta(t_1)) = 0 . \quad (3.23)$$

Dies bedeutet aber, da δ normalverteilt ist, daß

$$\Phi_\delta(t_1)-\Psi_\delta(t_1) = c' t_1+c_0$$

in der Umgebung von Null. Die Randbedingungen liefern erstens

$$0 = \Phi_\delta(0) - \Psi_\delta(0) = c_0$$

und zweitens

$$0 = \tilde{E}(\delta)-E(\delta) = \frac{\partial \Psi_\delta(t)}{\partial t_{1_i}}\Big|_{t_1=0} - \frac{\partial \Psi_\delta(t_1)}{\partial t_{1_i}}\Big|_{t_1=0}$$

$$= \frac{\partial c' t_1}{\partial t_{1_i}}\Big|_{t_1=0} = c_i \quad i=1,\ldots,n-m \ .$$

Also ist $\Psi_\delta(t_1) - \Psi_\delta(t_1) = 0$ in der Umgebung von Null. Zusammen mit (3.21) impliziert dies $\Psi_\xi(t_1) = \Psi_\xi(t_1)$ und schließlich wegen (3.20) auch $\Psi_\varepsilon(t_2) = \Psi_\varepsilon(t_2)$ jeweils in der Umgebung von Null. Wir fassen das Ergebnis zusammen in der

Bemerkung 3.27

In $\mathfrak{M}_2$ sind bei Identifizierbarkeit von Π die Funktionen $\Psi_\delta(t_1)$, $\Psi_\xi(t_1)$ und $\Psi_\varepsilon(t_2)$ in der Umgebung der Null identifizierbar. Insbesonders ist also die Kovarianzmatrix Ω identifizierbar.

3.2 Identifikation i.s.S. im funktionalen Fall

Wir ändern nun die Annahme über die Folge $(\xi(t))_{t\in\mathcal{T}}$.

(A') $(\xi(t))$, $t\in\mathcal{T}=\{1,\ldots,T\}$ sei eine Folge von (nichtstochastischen) Vektoren aus dem $\mathbb{R}^{n-m}$, mit der Eigenschaft, daß $rg(\xi(1),\ldots,\xi(T)) = n-m$ ist.

Während der Parameter t wegen der Annahmen (A) und (B) in Abschnitt 3.1 unterdrückt werden konnte, gewinnt er hier durch die Annahme (A') an Bedeutung. Für jedes t erhalten wir in (1.3) eine neue unbeobachtbare Variable $\xi(t)$.

Wir bezeichnen mit $\mathfrak{M}_3$ das Modell [(1.3),(A'),(B),(C)] und setzen $\Sigma = \{(\Lambda,\xi(1),\ldots,\xi(T),\Omega) \mid \mathfrak{M}_3\}$,

$$\mathcal{K} = \Sigma$$

$$\mathcal{F} = \{F_{z(1),\ldots,z(T)} \mid \mathfrak{M}_3\}$$

und $\kappa \overset{\mathcal{K}}{\sim} \tilde{\kappa} :\Leftrightarrow \kappa = \tilde{\kappa}$, $\kappa, \tilde{\kappa} \in \mathcal{K}$.

Damit erhalten wir

Satz 3.28

In $\mathfrak{M}_3$ ist κ $(\overset{\mathcal{K}}{\sim}, \overset{\mathcal{F}}{\sim})$-identifizierbar.

Beweis: Aus $F_{z(1),\ldots,z(T)} = \tilde{F}_{z(1),\ldots,z(T)}$ folgt, daß $\Omega = \tilde{\Omega}$ und $\xi(t) = Ez(t) = \tilde{E}z(t) = \tilde{\xi}(t)$ für alle $t \in \mathcal{T}$. Aus der Gleichheit von $\xi(t)$ und $\tilde{\xi}(t)$ folgt mit der Rangbedingung in (A'), daß $\Pi = \tilde{\Pi}$ ist. ▲

Es treten also bei verteilungsgestützter Identifizierbarkeit im funktionalen Fall keine Probleme auf, da durch die angenommene Kenntnis der Verteilungsfunktion von $z(t)$ die eigentlich unbeobachtbaren Variablen $\xi(t)$ unter den gegebenen Annahmen durch die ersten Momente zu beschreiben sind.

Gleiches gilt im funktionalen Fall auch für momentengestützte Identifizierbarkeit. Bevor wir zur Untersuchung der momentengestützten Identifizierbarkeit im strukturellen Fall übergehen, sei noch darauf hingewiesen, daß sich im Abschnitt über trajektoriengestützter Identifizierbarkeit noch zeigen wird, wo die Probleme des funktionalen Falls liegen.

4. Momentengestützte Identifizierbarkeit

In diesem Abschnitt wählen wir die Beobachtungsäquivalenz so, daß zwei Verteilungen identifiziert werden, wenn sie in ihren Momenten bis zu einer gewissen Ordnung r_0 übereinstimmen, d.h.

$$F \overset{\mathfrak{M}}{\sim} \tilde{F} :\Leftrightarrow M_r(F) = M_r(\tilde{F}) \quad , \; r \leq r_0 \quad .$$

Dabei kann r_0 beliebig groß werden, solange die entsprechenden absoluten Momente existieren. Wir erinnern daran, daß wir im Abschnitt 3.1 keinerlei Annahmen über die Existenz irgendwelcher Momente von z gemacht haben.

In Zusammenhang mit Satz 3.7 hatten wir angenommen, daß die ξ-Komponenten nicht exakt beobachtbar sind. Wir wollen hier, da uns Momente zur Verfügung stehen, untersuchen, welche Auswirkungen die Möglichkeit, einige Variablen exakt beobachten zu können, auf die Identifizierbarkeit hat.

Wir nehmen an, daß mindestens zweite Momente existieren und betrachten (1.3) in der Form:

$$y = \Pi x + \varepsilon - \Pi\delta \tag{4.1}$$

unter den Annahmen des Modells $\mathfrak{M}_1$.

Der Vektor x sei unterteilt in $(x_1', x_2')'$, wobei $x_2 = \xi_2$ sei, also $\delta = (\delta_1', 0')'$, und $\Pi = (\Pi_1, \Pi_2)$. Mithin gilt auch

$$y = \Pi_1 x_1 + \Pi_2 x_2 + \varepsilon - \Pi_1 \delta_1 \quad . \tag{4.2}$$

Postmultiplikation von (4.2) mit x_2' und anschließender Erwartungswertbildung liefert

$$E(yx_2') = \Pi_1 E(x_1 x_2') + \Pi_2 E(x_2 x_2') + E(\varepsilon x_2') - \Pi_1 E(\delta_1 x_2') \tag{4.3}$$

Da $x_2 = \xi_2$, ist wegen der Annahmen (B) und (C)

$$E(\varepsilon x_2') = 0 \text{ und } E(\delta_1 x_2') = 0 \quad .$$

Unter der Annahme, daß $[E(x_2 x_2')]^{-1}$ existiert, folgt aus (4.3)

$$\Pi_2 = E(yx_2')[E(x_2x_2')]^{-1} - \Pi_1 E(x_1x_2')[E(x_2x_2')]^{-1} \quad . \quad (4.4)$$

Dies eingesetzt in (4.2) ergibt nach Umordnung

$$y - E(yx_2')[E(x_2x_2')]^{-1}x_2 = \Pi_1\{x_1 - E(x_1x_2')[E(x_2x_2')]^{-1}x_2\} + \varepsilon - \Pi_1\delta_1$$

bzw.

$$\tilde{y} = \Pi_1\tilde{x}_1 + \varepsilon - \Pi_1\delta_1 \quad , \qquad (4.5)$$

wobei in allen Komponenten von $\tilde{x}_1 = x_1 - E(x_1x_2')[E(x_2x_2')]^{-1}x_2$ Fehler enthalten sind. Ist Π_1 identifizierbar, so ist wegen der Beobachtbarkeit der Momente in (4.4) auch Π_2 identifizierbar. Somit ist also (4.5) nur weiter von Interesse. Dies stimmt aber der Form nach mit (4.1) überein. Wir werden also weiter mit (1.3) bzw. (4.1) arbeiten und o.B.d.A. annehmen, daß keine Komponente von ξ exakt zu beobachten ist.

4.1 Grad-r-momentengestützte Identifizierbarkeit, r>2

Wir stützen uns in diesem Unterabschnitt auf die Information, die durch die Momente der Verteilung von z bis zu einer Ordnung $r > 2$ gegeben ist. Existieren alle absoluten Momente r-ter Ordnung von z, so läßt sich die c.F. von z als Taylorreihe entwickeln (Vogel[1970],S.232):

$$\phi_z(t) = 1 + \sum_{p=1}^{r} \frac{[(t'\mathrm{grad})^p\phi](0)}{p!} + o(\|t\|^r)$$

$$= 1 + \sum_{p=1}^{r} \sum_{p_1+\ldots+p_n=p} \frac{(it_1)^{p_1}\cdots(it_n)^{p_n}}{p_1!\cdots p_n!} m_z(p_1,\ldots,p_n) + o(\|t\|^r)$$

Für genügend kleine t-Werte können wir $\phi(t)$ für $1+s$ in der Reihenentwicklung

$$\log(1+s) = \sum_{q=1}^{\infty} \frac{(-1)^{q+1}s^q}{q}$$

einsetzen. Nach Umordnung und Zusammenfassung erhalten wir eine Reihe der Form

$$\log\phi_z(t) = \sum_{p=1}^{r} \sum_{p_1+..+p_n=p} \frac{(it_1)^{p_1}\cdots(it_n)^{p_n}}{p_1!\cdots p_n!} \kappa_z(p_1,\ldots,p_n) +$$

$$+ o(\|t\|^r),$$

wobei die Koeffizienten $\kappa_z(p_1,\ldots,p_n)$ per definitionem die Kumulanten von z sind.

Die Kumulanten haben gegenüber Momenten theoretische Vorteile. Ist z.B. $z = u + v$ und sind u und v stochastisch unabhängig, so gilt: $\kappa_z = \kappa_u + \kappa_v$. Außerdem ist bekannt, daß alle Kumulanten der Normalverteilung existieren und diejenigen der Ordnung größer als 2 Null sind.

Wir wollen nun eine Beziehung zwischen Kumulanten von z und den Koeffizienten in Π finden und gehen dazu von der (n-m+j)-ten Gleichung des Systems (1.3) aus:

$$y_j = \Pi_j.\xi + \varepsilon_j \quad .$$

$\zeta^{(j)}$ sei der Vektor $(\xi_1,\ldots,\xi_{n-m},\eta_j)'$ und mit $\kappa_{\zeta^{(j)}}(p_1,\ldots,p_{n-m},p_{n-m+j})$ werden die zugehörigen Kumulanten bezeichnet. Um die gewünschte Beziehung herzuleiten, können wir von formalen Potenzreihen ausgehen. Wegen $\zeta^{(j)} = \begin{pmatrix} I \\ \Pi_j. \end{pmatrix}\xi$ gilt $\log\phi_\xi[(I,\Pi_j.')t] = \log\phi_{\zeta^{(j)}}(t)$ oder ausführlich:

$$\sum_{p=1}^{\infty} \sum_{p_1+..+p_{n-m}=p} \frac{i^p \kappa_\xi(p_1,\ldots,p_{n-m})}{p_1!\cdots p_{n-m}!} \prod_{l=1}^{n-m} (t_l+\pi_{j,l}t_{n-m+j})^{p_l}$$

$$= \sum_{p} \sum_{p_1+..+p_{n-m}=p} \frac{i^p \kappa_\xi(p_1,\ldots,p_{n-m})}{p_1!\cdots p_{n-m}!} \sum_{q_1=0}^{p_1} \cdots \sum_{q_{n-m}=0}^{p_{n-m}} \prod_{l=1}^{n-m} \binom{p_l}{q_l} \times$$

$$\times \pi_{j,l}^{p_l-q_l} \cdot t_l^{q_l} \cdot t_{n-m+j}^{p_l-q_l}$$

$$= \sum_{q=1}^{\infty} \sum_{q_1+..+q_{n-m}+q_{n-m+j}=q} \frac{i^q \kappa_{\zeta^{(j)}}(q_1,\ldots,q_{n-m},q_{n-m+j})}{q_1!\cdots q_{n-m}!q_{n-m+j}!} \times$$

$$\times \prod_{l=1}^{n-m} t_l^{q_l} \cdot t_{n-m}^{q_{n-m+j}} \quad .$$

Durch Koeffizientenvergleich erhalten wir

$$\frac{\kappa_\zeta(j)(q_1,\ldots,q_{n-m},q_{n-m+j})}{q_1!\cdots q_{n-m}!\,q_{n-m+j}!} = \sum_{\substack{p_1+\ldots+p_{n-m}\\=q_1+\ldots+q_{n-m}+q_{n-m+j}\\p_\ell\geq q_\ell}} \frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{p_1!\cdots p_{n-m}!} \times$$

$$\times \prod_{l=1}^{n-m}\binom{p_l}{q_l}\,\pi_{j,l}^{\,p_l-q_l}$$

bzw.

$$\kappa_\zeta(j)(q_1,\ldots,q_{n-m},q_{n-m+j}) = q_{n-m+j}!\sum_{\substack{\text{wie}\\\text{oben}}}\frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{(p_1-q_1)!\cdots(p_{n-m}-q_{n-m})!}\times$$

$$\times \prod_l \pi_{j,l}^{\,p_l-q_l} \qquad (4.6)$$

Daraus folgt unter Benutzung des Kronecker-Symbols $\delta_{\cdot,\cdot}$

$$\sum_{k=1}^{n-m}\pi_{j,k}\kappa_\zeta(j)(q_1+\delta_{1,k},\ldots,q_{n-m}+\delta_{n-m,k},q_{n-m+j}) =$$

$$= \sum_{k=1}^{n-m} q_{n-m+j}! \sum_{\substack{p_1+\ldots+p_{n-m}\\=q_1+\ldots+q_{n-m}+q_{n-m+j}+1\\p_\ell\geq q_\ell+\delta_{\ell,k}}} \frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{(p_1-q_1-\delta_{1,k})!\cdots(p_{n-m}-q_{n-m}-\delta_{n-m,k})!} \times$$

$$\times \prod_l \pi_{j,l}^{\,p_l-q_l}$$

$$= q_{n-m+j}!\sum_k\sum_{\substack{\text{wie}\\\text{oben}}}(p_k-q_k)\frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{(p_1-q_1)!\cdots(p_{n-m}-q_{n-m})!}\prod_l \pi_{j,l}^{\,p_l-q_l}$$

(da für $p_\ell=q_\ell$, $p_k-q_k=0$)

$$= q_{n-m+j}!\sum_k\sum_{\substack{\text{wie}\\\text{oben}\\\text{aber}\\p_\ell\geq q_\ell}}(p_k-q_k)\frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{(p_1-q_1)!\cdots(p_{n-m}-q_{n-m})!}\prod_l \pi_{j,l}^{\,p_l-q_l}$$

$$= q_{n-m+j}!\sum_{\substack{\text{wie}\\\text{oben}}}\frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{(p_1-q_1)!\cdots(p_{n-m}-q_{n-m})!}\prod_l \pi_{j,l}^{\,p_l-q_l}\,(\sum_k p_k-q_k)$$

$$= q_{n-m+j}!\,(q_{n-m+j}+1)!\sum_{\substack{\text{wie}\\\text{oben}}}\frac{\kappa_\xi(p_1,\ldots,p_{n-m})}{(p_1-q_1)!\cdots(p_{n-m}-q_{n-m})!}\prod_l \pi_{j,l}^{\,p_l-q_l}$$

(wegen (4.6))

$$= \kappa_\zeta(j)(q_1,\ldots,q_{n-m},q_{n-m+j}+1)\,.$$

Können wir nun n-m verschiedene Tupel $(q_1,\ldots,q_{n-m},q_{n-m+j})$ wählen mit $q_1+\ldots+q_{n-m}+q_{n-m+j} > 2$, so daß die entstehende Matrix K(siehe nächste Seite) nichtsingulär ist, so lassen sich die π_{jk}, j=1,...,m;k=1,...,n-m aus

$$\sum_{k=1}^{n-m} \pi_{jk}\kappa_z(j)(q_1+\delta_{1,k},\ldots,q_{n-m}+\delta_{n-m,k},q_{n-m+j}) =$$

$$= \kappa_z(j)(q_1,\ldots,q_{n-m},q_{n-m+j}+1) \quad , \; j=1,\ldots,m$$

berechnen, da erstens $\kappa_z = \kappa_\zeta + \kappa_\nu$ und zweitens wegen $q_1+\ldots+q_{n-m}+q_{n-m+j} > 2$ die entsprechenden κ_ν gleich Null sind.

$$K = \begin{pmatrix} \kappa_\zeta(j)(q_1^1+1, q_2^1, \ldots, q_{n-m}^1, q_{n-m+j}^1) & \cdots\cdots & \kappa_\zeta(j)(q_1^1, \ldots, q_{n-m}^1+1, q_{n-m+j}^1) \\ \vdots & & \vdots \\ \kappa_\zeta(j)(q_1^{n-m}+1, q_2^{n-m}, \ldots, q_{n-m}^{n-m}, q_{n-m+j}^{n-m}) & \cdots\cdots & \kappa_\zeta(j)(q_1^{n-m}, \ldots, q_{n-m}^{n-m}+1, q_{n-m+j}^{n-m}) \end{pmatrix}$$

(Superskripte $\hat{=}$ Indizes)

4.2 Identifizierbarkeit i.w.S.

Viele Schätzverfahren arbeiten mit ersten und zweiten empirischen Momenten. Um diese anwenden zu können müssen zumindest die Parameter, die geschätzt werden sollen, aus den theoretischen Entsprechungen dieser Momente heraus identifizierbar sein. Diesem Problemkreis wollen wir uns nun widmen und definieren die Äquivalenzbeziehung $\overset{\mathfrak{M}}{\sim}$ auf $\mathfrak{F}$ jetzt als:

$$F_z \overset{\mathfrak{M}}{\sim} \tilde{F}_z :\Leftrightarrow M_r(F_z) = M_r(\tilde{F}_z) \ , \ r \leq 2 \ .$$

In diesem Fall wird das Identifikationsproblem zu einem algebraischen Problem, dessen (negative) Lösung innerhalb des Modells $\mathfrak{M}_1$ bereits aus der Bemerkung 3.10 folgt. Da Zufallsvariablen, die nur durch ihre ersten beiden Momente bekannt sind, von normalverteilten Zufallsvariablen mit den gleichen Momenten erster und zweiter Ordnung nicht zu unterscheiden sind, gilt:

In $\mathfrak{M}_1$ ist Λ nicht i.w.S. identifizierbar.

Obgleich wir dieses Resultat nun schon haben, wollen wir uns hierzu die Momentengleichungen in $\mathfrak{M}_1$ ansehen:

$$Ez = \Lambda E\xi$$
$$E(zz') = \Lambda E(\xi\xi')\Lambda' + \Omega$$

Unter Ausnutzung der speziellen Struktur des FV-Modells erhalten wir

$$Ey = \Pi Ex \tag{4.7}$$

und

$$E(xx') = E(\xi\xi') + \Omega_{\delta\delta} \tag{4.8}$$

$$E(yy') = \Pi E(\xi\xi')\Pi' + \Omega_{\varepsilon\varepsilon} \tag{4.9}$$

$$E(yx') = \Pi E(\xi\xi') + \Omega_{\varepsilon\delta} \ . \tag{4.10}$$

Da wir durch die Möglichkeit des Auftretens von Absolutgliedern in (1.3) o.B.d.A. die Annahme treffen können, daß $Ex = E\xi = 0$ ist, entfällt (4.7) als zusätzliche Information. Da in $\mathfrak{M}_1$ bezüglich $E(\xi\xi')$ keiner-

lei Restriktionen bekannt sind, wird (4.8) für gegebenes $\Omega_{\delta\delta}$ ganz dazu benutzt, um $E(\xi\xi')$ zu bestimmen. Gleiches gilt übertragen für $\Omega_{\varepsilon\varepsilon}$ und (4.9). Somit bleibt zur Identifikation von Π, $\Omega_{\delta\delta}$ und $\Omega_{\varepsilon\delta}$ (4.10) in der Form

$$E(yx') = \Pi[E(xx') - \Omega_{\delta\delta}] + \Omega_{\varepsilon\delta} \qquad (4.11)$$

übrig. Wie wir bereits wissen, reicht dies aber in $\mathfrak{M}_1$ nicht aus.

Das Modell $\mathfrak{M}_1$ ist also in diesem Fall sinnlos. Der Vollständigkeit halber wollen wir hier zwei Fälle angeben, die in der Literatur auftauchen und Identifizierbarkeit unter Umständen ermöglichen, deren Annahmen aber sehr stark sind.

Sind einmal $\Omega_{\delta\delta}$ und, falls $\Omega_{\varepsilon\delta} \neq 0$, auch $\Omega_{\varepsilon\delta}$ bekannt, so läßt sich aus (4.11) Π eindeutig bestimmen, sofern $E(xx') - \Omega_{\delta\delta}$ nichtsingulär ist.

Wäre alternativ Ω diagonal, d.h. kann man annehmen, daß alle Fehler untereinander unkorreliert sind, so besteht die Möglichkeit zur Identifizierbarkeit. Notwendig hierzu ist, daß die $n(n+1)/2$ Gleichungen aus $E(zz')$ zahlenmäßig die

$m(n-m) + (n-m)(n-m+1)/2 + n$ Unbekannten in Λ, $E(\xi\xi')$ und Ω übertreffen, d.h. daß

$$n > m \geq -(1-\sqrt{8n+1})/2$$

gilt. Diese Bedingung ist natürlich i.a. nicht hinreichend.

Auf der Suche nach besseren Auswegen aus der Nichtidentifizierbarkeit setzen wir die Betrachtung von (4.11) noch etwas fort. Diese Gleichung entsteht aus (4.1) durch Postmultiplikation mit x' und Erwartungswertbildung. Die störenden Terme $\Omega_{\delta\delta}$ und $\Omega_{\varepsilon\delta}$ tauchen auf, da in (4.1) die Fehler $\varepsilon - \Pi\delta$ mit dem Regressor x korreliert sind. Gäbe es nun andere beobachtbare Variablen, sagen wir u, mit

$$E[(\varepsilon - \Pi\delta)u'] = 0 \ , \qquad (4.12)$$

bestünde die Möglichkeit, Π zu identifizieren, vorausgesetzt,

$$E(xu') \text{ hat vollen Zeilenrang.} \qquad (4.13)$$

Denn (4.1) postmultipliziert mit u' ergibt

$$yu' = \Pi xu' + (\varepsilon - \Pi\delta)u' \ .$$

Erwartungswertbildung liefert

$$E(yu') = \Pi E(xu')$$

und diese Gleichung ist genau dann eindeutig lösbar, wenn (4.13) gilt und ein wahres Π^* existiert.

Variablen u, die (4.12) und (4.13) erfüllen, werden Instrumentenvariablen genannt. Sie wurden durch Reiersøl [1941] und Geary[1943] in die Literatur eingeführt und Sargan[1958] untersuchte weitgehend ihre Eigenschaften. Monfort[1978] benutzt den Instrumentenvariablengedanken, um Identifizierbarkeitsresultate in klassischen linearen Modellen herzuleiten.

Wir werden in den nächsten beiden Unterabschnitten zwei Fälle betrachten, in denen es die Modellannahmen ermöglichen, Instrumentenvariablen anzugeben.

4.2.1 Identifikation bei Nullrestriktionen

In der Praxis hat man oft Informationen, die sich auf die strukturellen Gleichungen des Systems (1.1-1.2) beziehen. Wir wollen hier den Fall von Nullrestriktionen betrachten, d.h. wir machen folgende Annahme:

(D) Die Parametermatrizen $B \in \mathbb{R}^{m\times(n-m)}$ und $\Gamma \in \mathbb{R}^{m\times m}$ in (1.1) sind bis auf folgende Einschränkungen vollständig unbekannt:

(i) Γ ist nichtsingulär
(ii) $\gamma_{jj} = -1 \ , \ j=1,\ldots,m$
(iii) den i-ten Zeilen $B_{i.}$ und $\Gamma_{i.}$ von B und Γ

sind Nullrestriktionen auferlegt, so daß sie in der folgenden Form geschrieben werden können:

$$B_{i.} = \underset{1\times(n-m)^{*}}{(\ B_{i.}^{*}\ ,} \ \underset{1\times(n-m)^{**}}{O\ )}$$

$$\Gamma_{i.} = \underset{1\times m^{\Delta}}{(\ \Gamma_{i.}^{\Delta}\ ,} \ \underset{1\times m^{\Delta\Delta}}{O\ )}$$

Die erste Gleichung z.B. des Systems (1.1-1.2) stellt sich dann folgendermaßen dar:

$$y_1 = \Gamma_{1.}^{\dagger} y^{\dagger} + B_{1.}^{*} x^{*} + \varepsilon_1^{(2)} - \varepsilon_1^{(1)} - \Gamma_{1.}^{\dagger} \varepsilon^{(2)\dagger} - B_{1.}^{*} \delta^{*} \ .$$

Dabei ist $\Gamma_{1.}^{\dagger}$ durch $\Gamma_{1.}^{\Delta} = (-1, \Gamma_{1.}^{\dagger})$ erklärt, $y^{\Delta} = \begin{pmatrix} y_1 \\ y^{\dagger} \end{pmatrix}$ ist der Vektor der in der ersten Gleichung zugelassenen endogenen Variablen und x^{*} der Vektor der in der ersten Gleichung zugelassenen exogenen Variablen. Wir fassen die obige Gleichung zusammen in

$$y_1 = A_{1.}^{\dagger} w^{\dagger} + \varepsilon_1^{\dagger} \ ,$$

wobei $A_{1.}^{\dagger} = (\Gamma_{1.}^{\dagger}, B_{1.}^{*})$, $w^{\dagger} = (y^{\dagger\prime}, x^{*\prime})'$ und $\varepsilon_1^{\dagger}$ den Rest umfaßt.

Als Instrumenvariablen für die erste Gleichung kommen nun in Betracht:

- zugelassene exakt beobachtbare exogene Variablen (zusammengefaßt im $(n-m)_1^{*} \times 1$-Vektor ξ^{*})
- ausgeschlossene exakt beobachtbare exogene Variablen (zusammengefaßt im $(n-m)_1^{**} \times 1$-Vektor ξ^{**})
- ausgeschlossene exogene Variablen mit Fehlern, die mit etwaigen Fehlern der zugelassenen endogenen und exogenen Variablen unkorreliert sind (zusammengefaßt im $(n-m)_2^{**} \times 1$-Vektor $\tilde{x}^{**}$).

Bei dieser Aufzählung gelten für die erste Gleichung auch solche fehlerbehafteten exogenen Variablen als exakt beobachtbar, deren Fehlervarianzen und -kovarianzen

mit anderen Fehlern aus den übrigen Gleichungen bekannt sind.

Wir fassen ξ^*, ξ^{**}, $\tilde{x}^{**}$ im Vektor u zusammen und erhalten

Satz 4.1

Das System (1.1-1.2) unterliege den Annahmen (A),(B), und (D). Die zusätzliche Information bezüglich der ersten Gleichung

$$y_1 = A_1^{\dagger}.w^{\dagger} + \varepsilon_1^{\dagger}$$

bestehe aus

$$E(\varepsilon_1^{\dagger} u') = 0 .$$

Dann ist $A_1^{\dagger}.$ i.w.S. identifizierbar genau dann, wenn

$$rg[E(w^{\dagger} u')] = m^{\Delta} + (n-m)^* - 1 . \blacktriangle$$

Aus dieser 'Rangbedingung' lassen sich nun einfache notwendige 'Abzählbedingungen' ableiten:
$E(w^{\dagger} u')$ ist eine $[m^{\Delta}+(n-m)^*-1]\times[(n-m)_1^*+(n-m)_1^{**}+(n-m)_2^{**}]$-Matrix und soll diese den Rang $m^{\Delta}+(n-m)^*-1$ haben, so muß

$$(n-m)_1^*+(n-m)_1^{**}+(n-m)_2^{**} \geq m^{\Delta}+(n-m)^*-1$$

bzw. wenn $(n-m)^* = (n-m)_1^* + (n-m)_2^*$, wobei $(n-m)_2^*$ die Anzahl der zugelassenen fehlerhaft beobachtbaren exogenen Variablen angibt,

$$(n-m)_1^{**}+(n-m)_2^{**} \geq m^{\Delta}+(n-m)_2^*-1$$

gelten. Wir erhalten also folgendes

Korollar 4.2

Sind über die Voraussetzungen in Satz 4.1 hinaus keine weiteren Informationen bekannt, die zur Identifikation von $A_1^{\dagger}.$ beitragen könnten, so ist zur Identifikation i.w.S. von $A_1^{\dagger}.$ notwendig, daß es mindestens $m^{\Delta}+(n-m)^*-1$ exakt beobachtbare oder mit den Fehlern der zugelassenen Variablen unkorrelierte exogene Variablen gibt oder,daß

es mindestens $m^{\Delta}+(n-m)^{*}_{2}-1$ ausgeschlossene exogene Variablen gibt, die mit Fehlern der zugelassenen Variablen unkorreliert sind.

Wir können weiterhin eine Beziehung zu der klassischen Abzählbedingung für simultane Systeme ohne Fehler-in-den-Variablen herstellen, wenn wir zusätzlich die Identifizierbarkeit von Varianzen und Kovarianzen der Fehler der exogenen Variablen in Betracht ziehen. Zur Identifikation
-erstens der Koeffizienten in A_1.
-zweitens der durch die erste Gleichung hervorgerufenen Kovarianzen und
-drittens der Varianzen der Fehler der exogenen Variablen
stehen (n-m) Momentengleichungen aus (4.11) zur Verfügung. Es werden $m^{\Delta}+(n-m)^{*}-1$ dieser n-m Gleichungen dazu benutzt, die unbekannten Koeffizienten zu bestimmen. Somit bleiben $n-m-(m^{\Delta}+(n-m)^{*}-1) = (n-m)^{**}-m^{\Delta}+1$ Momentengleichungen übrig, um die Varianzen und Kovarianzen festzulegen. Da man aber hier einerseits nicht mehr Unbekannte bestimmen kann, als man Gleichungen hat, andererseits zusätzliche Gleichungen nur vorhanden sind, wenn

$$(n-m)^{**} \geq m^{\Delta} - 1,$$

erhalten wir als notwendige Bedingung zur Identifizierbarkeit der Koeffizienten, Varianzen und Kovarianzen der ersten strukturellen Gleichung, daß diese im klassischen Fall überidentifiziert sein muß und zwar um so viel, wie unbekannte Varianzen und Kovarianzen vorhanden sind. Wir halten dies fest in einem weiteren

Korollar 4.3

Unter den Bedingungen von Korollar 4.2 ist zur Identifikation der Koeffizienten, Varianzen und Kovarianzen notwendig, daß diese Gleichung im klassischen Fall um so viel überidentifiziert wäre, wie unbekannte Varianzen und Kovarianzen zu bestimmen sind.

Zur Verdeutlichung der beiden Korollare geben wir ein Beispiel aus der Arbeit von Geraci[1974] an. Folgende Beziehungen mögen gelten:

$$-y_1 + \gamma_{12}y_2 = \beta_{11}\xi_1 + \beta_{14}\xi_4 + \varepsilon_1$$

$$\gamma_{21}y_1 - y_2 = \beta_{21}\xi_1 + \beta_{22}\xi_2 + \beta_{23}\xi_3 + \varepsilon_2$$

und $x_2 = \xi_2$, $x_3 = \xi_3$, $x_4 = \xi_4$.

Betrachten wir zunächst die erste Gleichung:

Die einzigen ausgeschlossenen exogenen Variablen sind die exakt beobachtbaren ξ_2 und ξ_3, d.h., es ist in jedem Fall

$$^{(1)}\xi^{**} = \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}, \text{ also } {}^{(1)}(n-m)_1^{**} = 2$$

und $$^{(1)}(n-m)_2^{**} = 0.$$

Weiter ist

$$^{(1)}\xi^{*} = \xi_4 , \quad \text{also } {}^{(1)}(n-m)_1{}^{*} = 1.$$

Denn während klar ist, daß ξ_4 zu ξ^* gehört, überlegt man sich, daß x_1 nicht dazu zu rechnen ist, da die Kovarianz $\omega_{\delta_1\varepsilon_1}$ nicht aus der zweiten Gleichung bekannt sein kann, da sie dort gar nicht auftritt.

Für die erste Gleichung gilt außerdem

$$^{(1)}m^{\Delta} = 2 \quad \text{und} \quad {}^{(1)}(n-m)^{*} = 2,$$

so daß die notwendige Bedingung aus Korollar 4.2 erfüllt ist, diejenige aus Korollar 4.3 hingegen nicht, da $^{(1)}(n-m)^{**} - {}^{(1)}m^{\Delta} + 1 = 1$, aber zwei unbekannte Varianzen und Kovarianzen zu bestimmen sind.

Entsprechende Überlegungen bezüglich der zweiten Gleichung ergeben

$$^{(2)}\xi^{*} = \begin{pmatrix}\xi_2\\ \xi_3\end{pmatrix}, \text{ also } {}^{(2)}(n-m)_1{}^{*} = 2,$$

$$^{(2)}\xi^{**} = \xi_4 , \quad \text{also } {}^{(2)}(n-m)_1^{**} = 1$$

und $$^{(2)}(n-m)_2^{**} = 0.$$

Es ist $^{(2)}m^{\Delta} = 2$ und $^{(2)}(n-m)^{*} = 3$, so daß die Bedingung von Korollar 4.2 nicht erfüllt ist.

Nähmen wir wie Geraci[1974],[1976] an, daß $E(\delta_1\varepsilon_1) = E(\delta_1\varepsilon_2) = 0$, würde sich das Bild wesentlich ändern. In diesem Fall wäre $\omega_{\delta_1\delta_1}$ aus der ersten Gleichung bestimmbar und somit bestünde $^{(1)}\xi^*$ aus ξ_2, ξ_3 und x_1. Hierdurch wäre auch für die zweite Gleichung die Bedingung aus Korollar 4.2 erfüllt.

Abschließend sei noch bemerkt, daß in diesem Beispiel die Bedingungen bis auf singuläre Fälle sogar hinreichend zur Identifizierbarkeit sind.

Geraci[1974],[1976] hat sich ausführlich mit Modellen der eben behandelten Art beschäftigt. Er benutzt den 'impliziten Funktionen Satz', um im wesentlichen notwendige Bedingungen zur lokalen Identifizierbarkeit des gesamten Modells angeben zu können. Hsiao[1976] erzielt mit der gleichen Methode unter etwas schwächeren Annahmen ähnliche Resultate wie wir.

4.2.2 Identifikation bei seriell korrelierten exogenen Variablen

Ein weiteres Beispiel für die Möglichkeit, Instrumentenvariablen angeben zu können, bietet der Fall zeitlich korrelierter Beobachtungen. Da dies etwas vom zentralen Modell dieser Arbeit abweicht, wollen wir nur kurz darauf eingehen. Um diese von Reiersøl[1941] R. Frisch zugeschriebene Idee demonstrieren zu können, müssen wir den Parameter t wieder einführen und fassen deshalb die Zufallsvektoren $y(t) = (y_1(t),\ldots,y_m(t))'$ zu verschiedenen Zeitpunkten $t = 1,\ldots,T$ in einem großen Vektor $\text{vec}Y = (y_1(1),\ldots,y_m(1),\ldots,y_1(T),\ldots,y_m(T))'$ zusammen und verfahren bei den anderen Vektoren ebenso.

Wir erhalten statt (4.1)

$$\text{vec}Y = (I\otimes\Pi)\text{vec}X + \text{vec}\in - (I\otimes\Pi)\text{vec}\Delta$$

und statt (4.11)

$$E[\text{vec}Y\cdot(\text{vec}X)'] = (I\otimes\Pi)E[\text{vec}X\cdot(\text{vec}X)'] + E[\text{vec}\in\cdot(\text{vec}X)'] - (I\otimes\Pi)E[\text{vec}\Delta\cdot(\text{vec}X)'] \quad (4.14)$$

Unter (B) und der weiteren Annahme, daß der Zufallsprozeß (z(t)) schwach stationär ist, sehen wir, daß Π 2(T-1) Gleichungssystemen genügen muß:

$$M_{yx}(\ell) = \Pi M_{xx}(\ell) \quad , \quad \ell=\pm1,\ldots,\pm(T-1), \qquad (4.15)$$

wobei $M_{yx}(\ell) = E[y(t)x'(t-\ell)]$ ist. Für festes ℓ erhalten wir jeweils eine Lösungsmannigfaltigkeit

$$\Pi_0^{(\ell)} + \text{Kern}M_{xx}(\ell),$$

wobei $\Pi_0^{(\ell)}$ eine spezielle Lösung des Systems (4.15) ist. Π ist eindeutig bestimmt, wenn

$$\bigcap_{\substack{\ell=1-T \\ \ell\neq 0}}^{T-1} \text{Kern}M_{xx}(\ell) = 0. \qquad (4.16)$$

Speziell ist das der Fall, wenn für mindestens ein ℓ, $M_{xx}(\ell)$ nichtsingulär ist. Dann läßt sich $x(t-\ell)$ als Instrumentenvektor interpretieren.

Wir wollen die Bedingung (4.16) noch interpretieren. Ist $M_{xx}(\ell)$ singulär, so gilt $\det M_{xx}(\ell) = 0$. Es gibt dann einen Vektor $a\neq 0$, mit

$$0=M_{xx}(\ell)a=E[x(t)x'(t-\ell)]a=E[x(t)x'(t-\ell)a],$$

d.h. es gibt eine Linearkombination der Komponenten von $x(t-\ell)$, die orthogonal zu den Komponenten von x(t) ist. Soll (4.16) gelten, so bedeutet dies, daß es für verschiedene Laglängen ℓ nur verschiedene Linearkombinationen geben darf, die orthogonal zu den Komponenten des heutigen x sind.

Diese Möglichkeit der Identifikation hängt selbstverständlich entscheidend an der Tatsache, daß die Fehler keine der Korrelationsstruktur der exogenen Variablen entsprechende Korrelationsstruktur aufweisen, sondern etwa, wie hier, seriell unkorreliert sind.

Bisher wurde eine zeitliche Abhängigkeit angenommen, die nicht weiter spezifiziert wurde. Man kann nun aber den Standpunkt vertreten, daß diese nur Ausdruck einer fundamentaleren zugrundeliegenden Struktur ist. Es bietet sich

z.B. an, für die exogenen Variablen AR(MA)- Prozesse anzusetzen. Solche und ähnliche Spezifikationen sind die Grundlage der neueren Arbeiten (Mehra[1975],Maravall[1979], Maravall&Aigner[1977],Hsiao[1977],[1979],Nowak[1976],[1977]) zu diesem Problemkreis.

5. Trajektoriengestützte Identifizierbarkeit und konsistente Schätzbarkeit: struktureller vs. funktionaler Fall

Wir wollen die Arbeit mit einigen Bemerkungen zur trajektoriengestützten Identifizierbarkeit und konsistenten Schätzbarkeit abschließen. Bis hierhin standen wahrscheinlichkeitstheoretische und algebraische Probleme im Mittelpunkt, da wir von der Kenntnis der Verteilungen bzw. ihren Momenten ausgegangen waren. Diese Kenntnis besitzt man in der Praxis wohl kaum. Dort beobachtet man nur einen Teil einer Trajektorie bzw. in noch zu vertretender Idealisierung eine ganze Trajektorie. Somit stellen sich in der Schätzsituation noch zusätzlich statistische Probleme ein, die je nach dem zugrundeliegenden Fall unterschiedlich groß sind.

In den Abschnitten 3 und 4 haben wir gesehen, daß im strukturellen Fall verteilungs- bzw. momentengestützte Identifizierbarkeit nicht immer möglich ist, während im funktionalen Fall diesbezüglich keine Schwierigkeiten auftreten. Bei trajektoriengestützter Identifizierbarkeit und konsistenter Schätzbarkeit ist dies umgekehrt.

Wir demonstrieren dies, indem wir uns zuerst parallel zu den Resultaten im 3. und 4. Abschnitt die trajektoriengestützte Identifizierbarkeit und konsistente Schätzbarkeit für den strukturellen Fall anschauen und dann auf den funktionalen Fall eingehen.

5.1 Trajektoriengestützte Identifizierbarkeit im strukturellen Fall

Das Konzept der trajektoriengestützten Identifizierbarkeit hat als Verbindung zwischen verteilungs- und momentengestützter Identifizierbarkeit einerseits und konsistenter Schätzbarkeit andererseits theoretische Bedeu-

tung. Trajektoriengestützte Identifizierbarkeit liegt nach Definition 2.7 in zwei Fällen vor(die allerdings nur scheinbar verschieden sind, wie aus der Bemerkung 2.8 hervorgeht). Die theoretische Bedeutung ergibt sich aus der zweiten Möglichkeit in Zusammenhang mit den Aussagen der Bemerkungen 2.8 und 2.10.Sie ermöglicht, das Schätzproblem in zwei Stufen zu zerlegen.

In der ersten Stufe untersucht man die Erreichbarkeit der Verteilung der beobachtbaren Variablen bzw. bestimmter Momente davon. Ist als zweite Stufe die betreffende Charakteristik verteilungs- bzw. momentengestützt identifizierbar, so sichert das wesentliche Resultat (theorem 5) in der Arbeit von Deistler&Seifert [1978], daß bei Erreichbarkeit der externen Information die Charakteristik fast überall stark konsistent schätzbar ist.

Die Erreichbarkeit der externen Information bildet im strukturellen Fall kein Problem, da hier die Beobachtungen Realisationen unabhängig identisch verteilter Zufallsvariablen sind und somit eine wesentliche Voraussetzung für die verschiedenen Gesetze der großen Zahlen erfüllt ist.

5.2 Konsistente Schätzbarkeit im strukturellen Fall

Die oben angestellten Überlegungen sichern im strukturellen Fall konsistente Schätzbarkeit fast überall. Tatsächlich kann man aber 'überall' konsistent schätzen. Die meisten der folgenden Schätzverfahren sind, im Gegensatz zu der oben beschriebenen Möglichkeit, vom direkten Typ, d.h. sie gehen nicht explizit über die externe Information, sondern schätzen die Charakteristik unmittelbar.

Unter der in 3.1.1 angegebenen Identifizierbarkeitsbedingung, daß keine Linearkombination der Komponenten von ξ normalverteilt ist, liefert die 'minimum-distance'-

Methode von Wolfowitz[1952],[1957] eine stark konsistente Schätzerfolge für $\Pi_{i.}$. Sie bestimmt den Schätzer als eine meßbare Funktion der beobachtbaren Variablen derart, daß die Verteilung der beobachtbaren Variablen (wenn der Schätzer anstelle der zu schätzenden Parameter und Verteilungen gesetzt wird) der empirischen Verteilung der beobachtbaren Variablen am 'nächsten' kommt. Diese Methode erlaubt auch im Reiersøl-Fall(n=2, m=1) die 'kanonischen' Vertreter der Beobachtungsäquivalenzklassen des Rests der Struktur stark konsistent zu schätzen. Dabei ist der kanonische Vertreter die Wahl von ϕ_ξ und ϕ_ν, bei der ϕ_ξ keinen normalen Faktor besitzt.

Für den Reiersøl-Fall zeigt Kiefer&Wolfowitz [1956], daß das Maximum-Likelihood-Verfahren unter der Nichtnormalitätsbedingung für ξ die Parameter π_0 und π_1 stark konsistent zu schätzen gestattet, wohingegen dies für den kanonischen Vertreter nicht zutrifft.

Neyman[1951] betrachtet einen etwas allgemeineren Fall, indem er eine Zerlegung der Fehler in einen normalen und nichtnormalen Teil zuläßt und konstruiert einen Schätzer, der bei Identifizierbarkeit schwach konsistent ist.

Beispiel eines zweistufigen Vorgehens wäre die Ausformulierung unseres Beweisgangs in 3.1.2 zu einem Schätzverfahren. Der Nachteil aller dieser Schätzverfahren ist, daß sie kaum praktisch verwendbar sind. Eine vollkommen andere Methode, die auch praktikabel ist, hat Wald[1940] vorgeschlagen. Sein Gruppierungsschätzer ist in identifizierbaren Situationen des Reiersøl-Falls unter gewissen Annahmen(im Beispiel auf Seite 11 etwa wären sie erfüllt) schwach konsistent.

Für den allgemeinen Fall schließlich, den wir in 3.1.3 betrachten, benutzt Jeeves[1954] die auf Seite 48 beschriebene Idee, um ein Schätzverfahren anzugeben, daß eine stark konsistente Schätzerfolge liefert, wenn $nn_2(z) = n-m$ ist.

Die in 4.1 hergestellte Beziehung zwischen Kumulanten und Koeffizienten bildet die Grundlage zu einer Va-

riation der im obigen Sinne zweistufigen Momentenmethode. Bei Nichtsingularität der dort auftretenden Matrix ist die Beziehung stetig, so daß ein Ersetzen der Kumulanten durch ihre empirische Approximationen einen schwach konsistenten Schätzer ergibt. Madansky[1959] weist auf Arbeiten von Scott[1950] und Drion[1951] hin, die statt Kumulanten Momente benutzen.

Zu den in Abschnitt 4.2 behandelten identifizierbaren Fällen existiert eine Fülle von alternativen Schätzansätzen(4.2.1: Jöreskog[1973],Browne[1974],Hsiao[1976], Hausmann[1977],Geraci[1977],Robinson[1977];4.2.2: Mehra [1975]), die entweder das ML-Prinzip anwenden oder Variationen der Kleinstquadrate-Methode darstellen und zu konsistenten Schätzern führen. Durch Jöreskog[1973] ist sogar ein funktionierender Algorithmus bereitgestellt worden.

5.3 Funktionaler Fall

Im funktionalen Fall erhält man bei jeder Beobachtung in Form der neuen wahren Variable einen weiteren unbekannten Parameter. Diese 'inzidentellen' Parameter verhindern, daß ohne zusätzliche Annahmen Information über die Verteilung der beobachtbaren Variablen aufgebaut werden kann. Dies kommt auch in der Likelihoodfunktion zum Ausdruck. Solari[1969] hat im Reiersøl-Fall mit unabhängigen Fehlern gezeigt, daß die Lösung der Likelihoodgleichung kein Maximum, sondern einen Sattelpunkt darstellt.

Eine Möglichkeit, diesem Problem zu begegnen, besteht darin, anzunehmen, daß die empirische Verteilung der inzidentellen Parameter asymptotisch gegen eine Verteilung strebt, deren Nichtnormalität ausreicht, um identifizieren zu können. Dies Vorgehen wählt Wolfowitz in der Arbeit [1953], wo er entsprechende Annahmen macht und die starke Konsistenz seines minimum-distance-Schätzers beweist.

Für weitere Möglichkeiten und eine Zusammenstellung von offenen Fragen in diesem Zusammenhang sei auf Moran

[1971] verwiesen. Ergänzend zu den dortigen Literaturangaben sind Dolby[1972],Dolby&Lipton[1972],Dolby&Freeman[1975],Dolby[1976] und Nussbaum[1976] zu nennen.

Literaturverzeichnis

Aigner,D.J. & A.S. Goldberger(eds)[1977]: Latent Variables in Socioeconomic Models, North-Holland Publ. Co., Amsterdam

Breiman,L.,L. LeCam & L. Schwartz[1964]: Consistent Estimates and Zero-One Sets, AMS 35, 157-161

Browne, M.W.[1974]: Generalized Least Squares Estimators in the Analysis of Covariance Structures, South African Statistical Journal 8, 1-24

Bunke,H. & O. Bunke[1974]: Identifiability and Estimability, Math. Operationsforsch. u. Statist. 5, 223-233

Cramer,H. & H. Wold[1936]: Some Theorems on Distribution Functions, J.London Math. Soc. 11, 290-294

Copas,J.B.[1972]: The Liklihood Surface in the Linear Functional Relationship Problem, JRSS(B) 34, 274-278

Cuppens,R.[1975]: Decomposition of Multivariate Probability, Academic Press, New York

Deistler,M. & H.G. Seifert[1978]: Identifiability and Consistent Estimability in Econometric Models, E'trica 46, 969-980

Dolby,G.R.[1972]: Generalized Least Squares and Maximum Liklihood Estimation of Non-linear Functional Relationships, JRSS(B) 34, 393-400

" [1976]: The Ultrastructural Relation: A Synthesis of the Functional and Structural Relations, Biometr. 63, 39-50

" & S. Lipton[1972]: Maximum Liklihood Estimation of the General Non-linear Functional Relationship with Replicated Observations and Correlated Errors, Biometr. 59, 121-129

" & T.G. Freeman[1975]: Functional Relationships Having Many Independent Variables and Errors with Multivariate Normal Distribution, J. Mult. Anal. 5, 466-479

Doob,J.[1948]: Application of the Theory of Martingales, Coll. Int. C.N.R.S.-Le Calcul des Probabilités et ses Applications, Lyon, 23-27

" [1953]: Stochastic Processes, Wiley&Sons, New York

Drion,E.F.[1951]: Estimation of the Parameters of a Straight Line and of Variances of the Variables, if they are Both Subject to Error, Indagationes Math. 13, 256-260

Duncan,O.D.[1975]: INtroduction to Structural Equation Models, Academic Press, New York

Frisch,R.[1934]: Statistical Confluence Analysis by Means of Complete Regression Systems, Univ. Økonomiste Institutt Publ. No. 5, Oslo

Gänssler,P.[1969]: Exakte Schätzer und konsistente Schätzerfolgen, ZfWverw.Geb. 12, 145-152

Geary,R.C.[1942]: Inherent Relations between Random Variables Proc. Royal Irish Acad. Ser.A 47, 63-76

" [1943]: Relations Between Statistics: The General and the Sampling Problem When the Examples are Large, Proc. Royal Irish Acad. Ser.A 49, 177-196

Geraci,V.J.[1976]: Identification of Simultaneous Equation Models with Measurement Error, J.Econometrics 4, 263-283

" [1977]: Estimation of Simultaneous Equation Models with Measurement Error, E'trica 45, 1243-1255

Goldberger,A.S.[1972]: Structural Equation Methods in the Social Sciences, E'trica 40, 979-1002

" & O.D. Duncan [1973]: Structural Equation Models in the Social Sciences, Seminar Press, New York

Griliches,Z.[1974]: Errors in Variables and Other Unobservables, E'trica 42, 971-998

Haavelmo,T.[1944]: The Probability Approach in Econometrics, E'trica 12, Suppl.

Hausman,J.A.[1977]: Errors in Variables in Simultaneous Equation Models, J.Econometrics 5, 385-401

Howrey,E.P.[1978]: The Use of Preliminary Data in Econometric Forecasting, Rev.Econ.&Statist. 60, 193-200

Hsiao,C.[1976]: Identification and Estimation of Simultaneous Equation Models with Measurement Error, IER 17, 319-339

" [1977]: Identification for a Linear Dynamic Simultaneous Error-Shock Model, IER 18, 181-194

Hurwicz,L[1950]: Generalization of the Concept of Identification, 245-257 in Koopmans,T.C.(ed.) [1950]

Jeeves,T.A.[1954]: Identification and Estimation of Linear Manifolds in n-Dimensions, AMS 25, 714-724

Jöreskog,K.G.[1973]: A General Method for Estimating a Linear Structural Equation System , Chapter 5, 85-112 in Goldberger, A.S.&Duncan,O.D.[1973]

Kagan,A.M., Y.V.Linnik, C.R.Rao[1973]: Characterization Problems in Mathematical Statistics, Wiley, New York

Kendall,M.G.&Stuart,A[1969]: The Advanced Theory of Statistics, Vol.1, Griffin, London

Kiefer,J&Wolfowitz,J.[1956]: Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters, AMS 27, 887-906

Koopmans,T.C.(ed.)[1952]: Statistical Inference in Dynamic Economic Models, Wiley&Sons, New York

" &Reiersol,O[1950]: The Identification of Structural Characteristics, AMS 21, 165-181

LeCam,L.[1958]: Les Propriétés Asymptotiques des Solutions de Bayes, Publ. l'Inst. Statist. de l'Univ. Paris, Vol.VII, 17-35

Lukacs,E.[1970]: Characteristic Functions, Griffin, London

Madansky,A.[1959]: The Fitting of Straight Lines when Both Variables are Subject to Error, JASA 54, 173-205

Malinvaud,E[1970]: Statistical Methods of Econometrics, North Holland Publ.Comp. Amsterdam

Maravall,A[1979]: Identification in Dynamic Shock-Error Models, Springer Lecture Notes, Springer Verlag, Berlin

" &Aigner, D.J.[1977]: Identification of the Dynamic Shock-Error Model, 349-363 in Aigner,D.J.&Goldberger,A.S.[1977]

Marschak,J[1950]: Statistical Inference in Economics. An Introduction, 1-50 in Koopmans T.C.(ed.)[1950]

Mehra,R.K.[1975]: Identification and Estimation of the Error-in-Variables-Model (EVM) in Structural Form, IIASA RM-75-38, Laxenburg

Moran,P.A.P.[1971]: Estimating Structural and Functional Relationships, J. of Mult .Anal.1, 232-255

Morgenstern,O[1965]: Über die Genauigkeit wirtschaftlicher Beobachtungen, Physica-Verlag, Wien-Würzburg

Neyman,J[1951]: Existence of Consistent Estimates of the Directional Parameter in a Linear Structural Relation between Two Variables, AMS 22, 497-512

Nowak,E[1976]: Identifikation Ökonometrischer Zeitreihen-Regressionen bei Fehlern in den Variablen, ohne Benutzung zusätzlicher Kenntnisse, Forschungsber. Inst. für Statist. und Wirtschaftstheorie. Univ. München, Serie Oe. Nr. 6

" [1977]: An Identification Method for Stochastic Models of Time Series Analysis with Error in the Variables, Forschungsber. ... Nr. 8a

Nussbaum,M.[1976]: Maximum Likelihood and Least Squares Estimation of Linear Functional Relationships, Math. Operationsforsch. u. Statist.7, 23-49

Rao,C.R.[1966]: Characterisation of the Distribution of Random Variables in Linear Structural Relations, Sankya 28, 251-260

" [1967]: On Vector Variables with a Linear Structure and a Characterisation of the Multivariate Normal Distribution, Bull. Inst. Int. Statist. 42, 1207-1213

" [1969]: A Decomposition Theorem for Vector Variables with a Linear Structure, AMS 40, 1845-1849

Reiersøl,O.[1941]: Confluence Analysis by Means of Lag Moments and other Methods of Confluence Analysis, E'trica 9, 1-23

" [1950]: Identifiability of a Linear Relation between Variables which are Subject to Error, E'trica 18, 375-389

Richmond,J.[1976]: Aggregation and Identification, IER 17, 47-56

Robinson,P.M.[1977]: The Estimation of a Multivariate Linear Relation , J. Multivar. Anal. 7, 409-423

Rothenberg,T.S.[1971]: Identification in Parametric Models, E'trica 39, 577-591

Sargan,J.D.[1958]: The Estimation of Economic Relationships Using Instrumental Variables, E'trica 26, 393-415

Schönfeld, P.[1975]: A Survey of Recent Concepts of Identification, Disc. Paper, Univ. Bonn

Scott,E.L.[1950]: Note on Consistent Estimates of the Linear Structural Relation Between Two Variables, AMS 21, 284-288

Solari,M.E.[1969]: The "Maximum Likelihood Solution" of the Problem of Estimating a Linear Structural Relationship JRSS(B) 31, 372-375

Sprent,P.[1970]: The Saddle Point of the Likelihood Surface for a Linear Functional Relationship, JRSS(B) 32, 432-434

Vogel,W.[1970]: Wahrscheinlichkeitstheorie, Vandenhoeck& Ruprecht, Göttingen

Wald,A[1940]: The Fitting of Straight Lines if Both Variables are Subject to Error, AMS 11, 284-300

Wolfowitz,J.[1952]: Consistent Estimators of the Parameters of a Linear Structural Relation, Skand. Aktuarietidskrift 35, 132-151

" [1953]: Estimation by the Minimum Distance Method, Ann. Inst. Statist. Mathem. 5, Tokyo, 9-23

" [1957]: The Minimum Distance Method, AMS 28, 75-88.

unveröffentlichte Dissertationen:

Geraci,V.J.[1974]: Simultaneous Equation Models with Measurement Error, SSRI Workshop Series 7407, Univ. of Wisconsin, Madison

Jeeves,T.A.[1952]: Identifiability and Almost-Sure Estimability of Linear Structures in n-Dimensions, Univ. of California

LEBENSLAUF

Am 6. April 1950 wurde ich in Long Branch im Staat New Jersey(N.J.), U.S.A. als zweites von drei Kindern der Eheleute Dr. Hans-Joachim und Anna(geb. Bickenbach) aufm Kampe geboren. Meine Familie wohnte bis zum 22. Mai 1962 in Colt's Neck, N.J., wo ich ab September 1955 die Atlantic(später Colt's Neck) Township School besuchte. Am 23. Mai 1962 siedelten wir in die Bundesrepublik Deutschland über und wohnten vom 1. Juni 1962 bis 7. Juni 1964 in Hunstig im Kreis Oberberg. Während dieser Zeit war ich Schüler am Städt. Gymnasium in Gummersbach. Am 8. Juni 1964 zog ich mit meinen Eltern und Geschwistern nach Rheinbreitbach, Kreis Neuwied. Hier verbrachte ich bis zu meinem Abitur am 10. Juni 1968 den Rest meiner Schulzeit am nahegelegenen Siebengebirgsgymnasium in Bad Honnef.

Das Studium der Mathematik begann ich am 2. September 1968 an der Rhein.-Fried.-Wilh.-Universität zu Bonn und legte am 16. Oktober 1970 das Vordiplom mit Nebenfach Physik ab. Am 23. April 1974 erwarb ich das Diplom in Mathematik mit Nebenfach Ökonometrie. Seit dem 1. Mai 1974 bin ich an der Abteilung Ökonometrie des Instituts für Ökonometrie und Operations Research der Uni Bonn beschäftigt, zunächst als Verwalter einer wissenschaftlichen Assistentenstelle, seit April 1977 als wissenschaftlicher Angestellter im Sonderforschungsbereich 21 der DFG.

Am 20. Februar 1976 heiratete ich Adelheid Walkembach aus Bad Honnef. Am 16. August 1977 wurde unsere Tochter Berit geboren. Vom 1. August 1976 bis 7. Juli 1977 wohnten wir in Bad Honnef, seitdem wohnen wir in Königswinter.